

# Srovnání modifikací predikčních bankrotních modelů

Diplomová práce

Vedoucí práce:

Ing. Vladimír Hajko, Ph.D.

Bc. Ondřej Bednář

Brno 2017



**Poděkování**

Chtěl bych poděkovat vedoucímu mé diplomové práce Ing. Vladimíru Hajkovi, Ph.D., za jeho cenné rady, názory a profesionální připomínky, díky kterým byla tato práce úspěšně dokončena.



## Čestné prohlášení

Prohlašuji, že jsem tuto práci: **Srovnání modifikací predikčních bankrotních modelů** vypracoval samostatně a veškeré použité prameny a informace jsou uvedeny v seznamu použité literatury. Souhlasím, aby moje práce byla zveřejněna v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů, a v souladu s platnou *Směrnicí o zveřejňování vysokoškolských závěrečných prací*.

Jsem si vědom, že se na moji práci vztahuje zákon č. 121/2000 Sb., autorský zákon, a že Mendelova univerzita v Brně má právo na uzavření licenční smlouvy a užití této práce jako školního díla podle § 60 odst. 1 Autorského zákona.

Dále se zavazuji, že před sepsáním licenční smlouvy o využití díla jinou osobou (subjektem) si vyžádám písemné stanovisko univerzity o tom, že předmetná licenční smlouva není v rozporu s oprávněnými zájmy univerzity, a zavazuji se uhradit případný příspěvek na úhradu nákladů spojených se vznikem díla, a to až do jejich skutečné výše.

V Brně dne 20. května 2017

---



## **Abstract**

Bednář, O. Comparison of the modifications of bankruptcy prediction models. Diploma theses. Brno: Mendel university, 2017.

The goal of this theses is to compare existing bankruptcy prediction models with its new modification unique for this work, which could perform better than its competition. Proposed model is logit-based and consists of the combination of variables used in Altman's and Ohlson's models. The final model is estimated for medium sized companies in EU which aren't publicly traded. This model achieved prediction accuracy of 97,1% (97.4% for healthy and 91.1% for bankrupt companies) on its original dataset. As expected, when verified on new dataset, the accuracy dropped but still reaches 97.1% (99.3% for healthy and 37.7% for bankrupt companies). The model is compared with its competition (original and modified version of Ohlson's and partially Altman's models) and it is shown that it has higher prediction accuracy.

## **Keywords**

Bankruptcy, prediction model, bankruptcy model, Altman, Ohlson, o-score, z-score, logit, hold-out sample, ROC analysis, AUC.

## **Abstrakt**

Bednář, O. Srovnání modifikací predikčních bankrotních modelů. Diplomová práce. Brno: Mendelova univerzita v Brně, 2017.

Cílem této práce je porovnat stávající bankrotní modely s jejich novou modifikací, jenž je unikátní pro tuhle práci a mohla by být úspěšnější než její v praxi využívané protějšky. Navrhovaný model je založen na formě logit a jeho obsah se skládá z kombinace vysvětlujících proměnných obsažených v původním Ohlsonově a Altmanově modelu. Finální model je zaměřen na středně veliké společnosti z oblasti EU, které nejsou obchodovány na akciových trzích. Tento model na původních datech dosahuje úspěšnosti predikce 97,1 % (97,4 % pro zdravé a 91,1 % pro bankrotní podniky). Při jeho ověření na verifikačním data setu se podle očekávání úspěšnost sníží, ale i tak dosahuje predikční úspěšnosti 97,1 % (99,3 % pro zdravé a 37,7 % pro bankrotní podniky). Na těchto datech je nový model porovnán s původními a modifikovanými verzemi Ohlsonova i Altmanova (pouze částečně) modelu. V tomto porovnání vychází model jako nejkvalitnější s nejvyšší úspěšností predikce.

## **Klíčová slova**

Bankrot, predikční model, bankrotní model, Altman, Ohlson, o-score, z-score, logit, verifikační vzorek, ROC analýza, AUC.





# Obsah

<b>1</b>	<b>Úvod</b>	<b>17</b>
1.1	Cíl práce a výzkumné otázky.....	17
<b>2</b>	<b>Literární přehled</b>	<b>18</b>
2.1	Vývoj bankrotních modelů .....	18
2.1.1	Účetní bankrotní modely.....	18
2.1.2	Tržní bankrotní modely.....	19
2.1.3	Kombinované bankrotní modely .....	19
2.1.4	Modifikace známých modelů vs. tvorba nových modelů.....	20
2.2	Hlavní typy bankrotních modelů .....	21
2.3	Modifikace O-score a Z-score ve světové literatuře.....	26
2.4	Modifikace vybraných modelů v české literatuře.....	29
<b>3</b>	<b>Výběr dat a proměnných</b>	<b>31</b>
3.1	Kritéria pro výběr společností .....	31
3.2	Výběr vysvětlujících proměnných .....	33
3.2.1	Odhad vlivu vysvětlujících proměnných.....	34
3.3	Úprava dat.....	38
<b>4</b>	<b>Metodika</b>	<b>40</b>
4.1	Logistická regrese .....	40
4.2	Metody hodnocení výsledných modelů .....	42
4.2.1	Klasifikační tabulka .....	42
4.2.2	ROC křivka .....	42
4.3	Spojitost s výzkumnými otázkami .....	44
<b>5</b>	<b>Výsledky</b>	<b>46</b>
5.1	Výběr dat a tvorba data setu.....	46
5.2	Modely pro české podniky (CZ) .....	48
5.2.1	Model CZ_1.1_výchozí model.....	48
5.2.2	Model CZ_2.1_dílčí modely.....	50

---

5.2.3	Model CZ_2.2_upravený model .....	51
5.2.4	ROC analýza původních dat.....	54
5.2.5	Verifikace modelu CZ .....	57
5.2.6	Souhrn modelu CZ.....	60
5.3	Modely pro data podniků z EU .....	61
5.3.1	Model EU_1.1_výchozí.....	62
5.3.2	Model EU_2.1_finální.....	64
5.3.3	Model EU_3.1_upravený.....	65
5.3.4	ROC analýza na původních datech.....	67
5.3.5	Verifikace modelu EU.....	73
5.4	Srovnání se známými modely .....	82
5.4.1	Aplikace: Ohlson .....	82
5.4.2	Aplikace: Altman.....	84
5.4.3	Porovnání modelů.....	86
5.5	Shrnutí výsledků .....	87
5.6	Omezení a doporučení.....	89
<b>6</b>	<b>Závěr</b>	<b>91</b>
<b>7</b>	<b>Literatura</b>	<b>93</b>
<b>A</b>	<b>Literární přehled</b>	<b>99</b>
<b>B</b>	<b>Vyhledávací kritéria pro databázi Amadeus</b>	<b>101</b>
<b>C</b>	<b>Model CZ</b>	<b>103</b>
<b>D</b>	<b>Model EU</b>	<b>105</b>

## Seznam obrázků

Obr. 1	Porovnání rozdělení logit a probit (převzato z Heij et al., 2004)	41
Obr. 2	ROC: dokonalý klasifikátor	43
Obr. 3	ROC: model nevhodný ke klasifikaci	44
Obr. 4	ROC: CZ_2.2_finální model	55
Obr. 5	S+S: CZ_2.2_finální model	55
Obr. 6	ROC: CZ_2.1_výchozí model	56
Obr. 7	ROC: CZ_Verifikace_upravený model	58
Obr. 8	ROC_CZ_Verifikace_finální model	59
Obr. 9	ROC: Model EU 3.2-3	68
Obr. 10	ROC: Model EU_2.2-3	69
Obr. 11	ROC: Model EU_1.2-3	70
Obr. 12	ROC: Model EU_3.2-2	71
Obr. 13	ROC: Model EU_2.2-2	72
Obr. 14	ROC: Model EU_1.2-2	72
Obr. 15	ROC: Verifikace 1-3	75
Obr. 16	ROC: Verifikace 2-3	75
Obr. 17	ROC: Verifikace 3-3	76
Obr. 18	ROC: Verifikace 1-2	77
Obr. 19	ROC: Verifikace 2-2	78
Obr. 20	ROC: Verifikace 3-2	79
Obr. 21	Rozložení dat: EU_1-2_souhrnné	80
Obr. 22	Rozložení dat: EU_1-2_bankrotní	81

<b>Obr. 23</b>	<b>ROC: Ohlson_1980</b>	<b>83</b>
<b>Obr. 24</b>	<b>ROC: Ohlson_2003</b>	<b>84</b>
<b>Obr. 25</b>	<b>ROC: Altman_1968</b>	<b>85</b>
<b>Obr. 26</b>	<b>ROC: Altman_1995</b>	<b>86</b>

## Seznam tabulek

<b>Tab. 1</b>	<b>Poměrové ukazatele použité Beaverem (1966)</b>	<b>22</b>
<b>Tab. 2</b>	<b>Poměrové ukazatele použité Altmanem (1968)</b>	<b>22</b>
<b>Tab. 3</b>	<b>Zóny Altmanova Z-score (1968)</b>	<b>23</b>
<b>Tab. 4</b>	<b>Poměrové ukazatele použité Ohlsonem (1980)</b>	<b>24</b>
<b>Tab. 5</b>	<b>Poměrové ukazatele použité Zmijewskim (1984)</b>	<b>25</b>
<b>Tab. 6</b>	<b>Vybrané vysvětlující účetní proměnné</b>	<b>34</b>
<b>Tab. 7</b>	<b>Odhadnuté vlivy a znaménka vysvětlujících proměnných</b>	<b>37</b>
<b>Tab. 8</b>	<b>Obecný tvar klasifikační tabulky</b>	<b>42</b>
<b>Tab. 9</b>	<b>Amadeus vyhledávání: Data set CZ_2013 aktivní</b>	<b>46</b>
<b>Tab. 10</b>	<b>Amadeus vyhledávání: Data set CZ_2013 bankrotní</b>	<b>47</b>
<b>Tab. 11</b>	<b>Model CZ_1.1</b>	<b>49</b>
<b>Tab. 12</b>	<b>Model CZ_1.2_výchozí model</b>	<b>49</b>
<b>Tab. 13</b>	<b>Model CZ_2.1_dílčí modely</b>	<b>50</b>
<b>Tab. 14</b>	<b>Model CZ_2.2_úprava <math>X_{2+7}</math></b>	<b>51</b>
<b>Tab. 15</b>	<b>Model CZ_2.2_úprava <math>X_{2+7+10}</math></b>	<b>52</b>
<b>Tab. 16</b>	<b>Model_CZ_2.2_úprava <math>X_{2+7+10+6}</math></b>	<b>52</b>
<b>Tab. 17</b>	<b>Model_CZ_2.2_všechny úpravy</b>	<b>53</b>
<b>Tab. 18</b>	<b>Model_CZ_2.2_finální model</b>	<b>53</b>
<b>Tab. 19</b>	<b>Úspěšnost modelu CZ_2.2_finální při <math>P = 0,1</math></b>	<b>56</b>
<b>Tab. 20</b>	<b>Úspěšnost modelu CZ_2.1 při <math>P = 0,1</math></b>	<b>57</b>
<b>Tab. 21</b>	<b>Úspěšnost modelu CZ_2.1 při <math>P = 0,665</math></b>	<b>57</b>
<b>Tab. 22</b>	<b>První verifikace modelu CZ při <math>P = 0,1</math></b>	<b>58</b>

<b>Tab. 23</b>	<b>Úspěšnost modelu CZ na nových datech při <math>P=0,665</math></b>	<b>59</b>
<b>Tab. 24</b>	<b>Použité země EU</b>	<b>62</b>
<b>Tab. 25</b>	<b>Použité země EU (upraveno)</b>	<b>63</b>
<b>Tab. 26</b>	<b>Model EU_1.1</b>	<b>63</b>
<b>Tab. 27</b>	<b>Model EU_1.2_výchozí</b>	<b>63</b>
<b>Tab. 28</b>	<b>Model EU_2.1</b>	<b>64</b>
<b>Tab. 29</b>	<b>Model EU_2.2_finální</b>	<b>64</b>
<b>Tab. 30</b>	<b>Model EU_3.1</b>	<b>66</b>
<b>Tab. 31</b>	<b>Model EU_3.2_upravený</b>	<b>66</b>
<b>Tab. 32</b>	<b>Model EU_3.2-3, pro <math>P=0,108</math> a <math>0,498</math></b>	<b>68</b>
<b>Tab. 33</b>	<b>Model EU_2.2-3, pro <math>P= 0,108</math> a <math>0,498</math></b>	<b>69</b>
<b>Tab. 34</b>	<b>Model EU_1.2-3, pro <math>P=0,108</math> a <math>0,498</math></b>	<b>70</b>
<b>Tab. 35</b>	<b>Model EU_3.2-2, pro <math>P=0,092</math> a <math>0,5</math></b>	<b>71</b>
<b>Tab. 36</b>	<b>Model EU_2.2-2, pro <math>P=0,092</math> a <math>0,5</math></b>	<b>71</b>
<b>Tab. 37</b>	<b>Model EU_1.2-2, při <math>P=0,092</math> a <math>0,5</math></b>	<b>73</b>
<b>Tab. 38</b>	<b>Evropské země k verifikaci modelu EU (očištěno)</b>	<b>74</b>
<b>Tab. 39</b>	<b>Úspěšnost verifikace 1-3</b>	<b>74</b>
<b>Tab. 40</b>	<b>Úspěšnost verifikace 2-3</b>	<b>76</b>
<b>Tab. 41</b>	<b>Úspěšnost verifikace 3-3</b>	<b>77</b>
<b>Tab. 42</b>	<b>Úspěšnost verifikace 1-2</b>	<b>78</b>
<b>Tab. 43</b>	<b>Úspěšnost verifikace 2-2</b>	<b>79</b>
<b>Tab. 44</b>	<b>Úspěšnost verifikace 3-2</b>	<b>80</b>
<b>Tab. 45</b>	<b>Úspěšnost predikce Ohlson 1980</b>	<b>83</b>
<b>Tab. 46</b>	<b>Úspěšnost predikce Ohlson 2003</b>	<b>84</b>

---

<b>Tab. 47</b>	<b>Porovnání modelů při aplikaci na verifikační data set</b>	<b>86</b>
<b>Tab. 48</b>	<b>Porovnání složení vytvořených a používaných modelů</b>	<b>89</b>
<b>Tab. 49</b>	<b>Altmanovo Z-metrics (Dluhošová, 2011)</b>	<b>99</b>
<b>Tab. 50</b>	<b>Souhrnný přehled současných modifikací O-score</b>	<b>100</b>
<b>Tab. 51</b>	<b>Amadeus EU 2013 výsledky</b>	<b>101</b>
<b>Tab. 52</b>	<b>Amadeus EU 2014 výsledky</b>	<b>102</b>





# 1 Úvod

Bankrot společnosti je dlouhodobě zkoumaným tématem, které se týká téměř všech ekonomik a zemí na světě. Když společnost zbankrotuje, tak negativně ovlivňuje nejen vlastníky, ale i investory, dodavatele a další entity, které ztrácí alespoň část hodnoty vložené do společnosti samotné nebo do spolupráce s ní. Proto se jej snaží účetní, akademici i jiné zainteresované strany pochopit a predikovat již několik desítek let.

Predikce bankrotu je široké téma, na které existuje nespočet prací, článků a diskuzí, protože je více způsobů, jak ji chápat a pokusit se o její provedení. Nejstarší způsoby se zabývaly jeho souvislostí s hodnotami poměrových ukazatelů získaných z účetnictví daných podniků. Postupem času, byly vytvořeny různé matematické modely, které tyto ukazatele chápou jako vysvětlující proměnné a bankrot se snaží vysvětlit pomocí jedné zobecněné hodnoty. Vývoj těchto modelů, ale poukazuje na skutečnost, že účetní data promítnuté do poměrových ukazatelů nestačí na kvalitní predikci bankrotu. Proto se v současnosti značná část expertů snaží vyvinout modely, které zahrnují i další proměnné, které nemají počátek v účetnictví, ale v jiných oblastech, jako makroekonomické indikátory nebo proměnné popisující trh, na kterých se společnost pohybuje.

Je zřejmé, že pro predikci bankrotu společností neexistuje jednotný dokonalý recept, ale je více přístupů s různou přesností odhadu v závislosti na podmínkách daného vzorku společností. Tato práce se zabývá shrnutím dosavadních zkušeností z této problematiky, provedením empirické komparace dominantních existujících specifikací a aplikováním vybraných metod predikce na konkrétním vzorku společností z prostředí České republiky.

## 1.1 Cíl práce a výzkumné otázky

Cílem práce je empirické srovnání existujících modifikací predikčních bankrotních modelů typu Ohlsonova o-score s vlastní specifikací a odhadem tohoto typu modelu na datech podniků v ČR. Jinými slovy se tato práce zabývá tvorbou predikčního bankrotního modelu. Tento predikční model je tvořen na základě dat získaných ze společností na území ČR a predikční schopnost jeho finální podoby je na konci práce porovnána s vybranými, dnes využívanými predikčními modely.

Jako součást této práce jsou definovány dvě výzkumné otázky, na které je v průběhu této práce hledána odpověď:

- Má finální predikční model vytvořený v této práci vyšší predikční sílu než náhodné hádání?
- Je finální navržený predikční model úspěšnější v predikci bankrotu než původní a aktualizované verze Ohlsonova a Altmanova modelu?

Blíže je postup pro získání odpovědí na tyto otázky obsažen v části popisující metodiku práce.

## 2 Literární přehled

Téma bankrotních modelů je velmi obsáhlé, a ne vždy je jednoduché se v současné literatuře zorientovat. Bankrotní modely často nabývají různých forem, proto se tato kapitola zabývá stručným shrnutím názorů, podmínek a autorů, kteří přispěli k vytvoření důležitých typů těchto modelů. Díky jejich pozdějšímu využití v této práci, jsou zde zmíněny i modifikace predikčních modelů známých autorů, ať už ve světové nebo domácí literatuře.

### 2.1 Vývoj bankrotních modelů

Tato podkapitola se zabývá stručným shrnutím historického vývoje v oblasti tvorby modelů pro predikci bankrotu. Už dříve bylo zmíněno, že na tvorbu bankrotních modelů je možno pohlížet z více úhlů, a proto zde budou ty nejdůležitější pomyslné směry uvedeny a popsány.

#### 2.1.1 Účetní bankrotní modely

Počátky predikce bankrotu se datují do 40. let dvacátého století, kdy se Fitzpatrick (1931) snažil porovnat zdravé a zbankrotované společnosti na základě analýzy poměrových ukazatelů. Využil 13 ukazatelů, které měly indikovat bankrot, ale model se neprokázal jako signifikantní ve vztahu k bankrotu, a proto není v odborné literatuře často zmiňován.

Beaver (1966, 1968) navázal na práci Fitzpatricka a také se pokusil o predikci bankrotu za použití poměrových ukazatelů, přičemž jeho práce je všeobecně považována za první významnou práci v této oblasti. Beaverovy ukazatele byly vypočítány převážně na základě dat získaných z účetních výkazů společností, a tím dal nevědomky vzniknout pomyslnému směru v oblasti predikce bankrotu, který se jej snažil vysvětlit pomocí účetních dat. Modelům, které operují na bázi účetních dat, se dodnes v odborné literatuře přezdívá účetní modely, a proto toto označení bude použito i dále v této práci.

Do tohoto pomyslného směru se později zařadil i Altman (1968), který pomocí metody diskriminační analýzy zjednodušil posuzování finančního zdraví podniku, a tedy i pravděpodobnosti jeho bankrotu. Toto zjednodušení spočívalo ve výpočtu jedné zobecněné hodnoty  $Z$ , která po přiřazení do patřičného intervalu, vypovídala o finančním zdraví podniku. Na základě této  $Z$  hodnoty se daný model označuje jako Altmanovo  $Z$ -score. Dalšími významnými osobnostmi, které se zabývali tvorbou účetních bankrotních modelů, je i Ohlson (1980) a Zmijewski (1984).

Ohlson ve své práci využil logit model, který finálním uživatelům zjednodušil interpretaci, protože výsledkem tohoto modelu je hodnota, která přímo určuje procentní pravděpodobnost bankrotu dané společnosti. Oproti Altmanovu  $Z$ -score tedy není zapotřebí intervalů, které vysvětlují vypočítanou hodnotu. Stejně jako u

Z-score je tento model označován po jeho finální hodnotě, tedy jako Ohlsonovo O-score.

Zmijewski (1984) se stejně jako Altman a Ohlson zaměřil ve své práci na predikci bankrotu pomocí poměrových ukazatelů. Stejně jako Ohlson se snažil změnit formu modelu, ale na rozdíl od Ohlsona, Zmijewski použil *probit* model se třemi vysvětlujícími proměnnými.

### 2.1.2 Tržní bankrotní modely

V předchozích modelech jsou základem pro predikci budoucí finanční tísně nebo bankrotu použity tradiční poměrové ukazatele získané z účetních výkazů. Shumway (2001) ale zjistil, že polovina proměnných (poměrových ukazatelů) v modelech Altmana (1968) a Zmijewskiho (1984) jsou pro predikci bankrotu statisticky nevýznamné. Naopak poukazuje na to, že některé tržní proměnné, jako je např. velikost podniku nebo historická návratnost na akci, jsou s bankrotem provázány velmi silně. Ve své práci Shumway zkomboval dvě účetní a tři tržní proměnné, což mělo za výsledek přesnější predikci než u čistě účetních modelů. Shumwayovi závěry byly podpořeny prací Chavy a Jarrova (2004), kteří do svého tržního modelu přidávali účetní proměnné, ale dosáhli tím pouze minimálního zlepšení v přesnosti predikce. Na druhou stranu Beaver et al. (2005) použitím podobného modelu, který vztáhli na delší období, znovu potvrdili účinnost účetních proměnných. Jejich výsledky poukazují na to, že účetní a tržní proměnné se doplňují a použití tržních proměnných snižuje predikční schopnosti účetních proměnných pouze v kratších časových úsecích. Shumway tedy svým způsobem otevřel cestu novému směru v predikci bankrotu, kterým je použití tržních proměnných nebo jejich kombinaci s účetními proměnnými.

I když jsou statistické modely hojně používány v praxi, tak v oblasti využití tržních proměnných se do popředí dostávají modely založené na teorii oceňování opcí. V současné literatuře několik autorů použilo Black-Scholesův model oceňování opcí pro predikci rizika bankrotu. Příkladem mohou být práce Crosbie a Bohn (2003), Hillegeist et al. (2004), Xu a Zhang (2009) nebo Vassalou a Xing (2004). Vassalou a Xing i Hillegeist et al. ve svých pracích potvrzují, že modely založené na teorii oceňování opcí jsou přesnější než klasické statistické modely, ale čistě tržní BSM (Black-Scholes-Merton) modely také nejsou v predikci bankrotu dokonalé.

### 2.1.3 Kombinované bankrotní modely

Hillegeist et al. (2004) vytvořili zajímavou práci, kde porovnávali přepočítané verze Ohlsona a Altmana oproti tržnímu BSM modelu. Výsledkem byla lepší predikce pomocí BSM modelu, ale důležitější byly argumenty, kterými zkritizovali oba typy modelů (tržní i účetní). Uvádí, že problémem účetních modelů jsou hlavně podhodnocená aktiva (vzhledem k jejich tržní ceně) a chybějící proměnná, která by ukazovala volatilitu aktiv. Na druhou stranu kritizují i čistě tržní BSM model, a to hlavně kvůli zjednodušeným předpokladům a faktu, že ne všechny informace mohou být obsaženy v ceně. Na tyto závěry navazovala práce Xu a Zhanga (2009), ve

kteřé autoři sestavili kombinovaný model, který využil jak účetních, tak tržních proměnných a vykazoval lepší predikční schopnost než jeho jednotlivé části.

Stejní autoři na příkladu japonské ekonomiky také ukazují, že na základě situace je vhodné do modelu zahrnout i jiné než tržní a účetní proměnné. Tyto specifické proměnné odrážejí určité unikátní vlastnosti daného odvětví nebo země. V tomto případě se jednalo o závislost společnosti na Keiretsu, což je skupina společností (většinou podpořená významnou bankou), které si navzájem finančně pomáhají. V jejich práci se předpokládalo, že v rámci Keiretsu budou společnosti více odolné vůči bankrotu než společnosti bez účasti v Keiretsu. Ve finálním modelu byly zahrnuty účetní, tržní i tyto specifické proměnné a daný kombinovaný model měl vyšší vypovídací schopnost než výše zmíněný model, který kombinoval „pouze“ tržní a účetní proměnné.

Z těchto závěrů je jasné, že čistě účetní modely jsou nedostačující a vykazují lepší výsledky, pokud jsou v kombinaci s proměnnými, které odrážejí i tržní nebo specifickou situaci dané oblasti.

#### 2.1.4 Modifikace známých modelů vs. tvorba nových modelů

Doposud bylo zmíněno pouze pomyslné dělení bankrotních modelů na tržní a účetní (kombinované) v závislosti na použitých proměnných, ale v současné době se akademici a praktici v této oblasti dělí i na dva další pomyslné směry. Vzhledem k faktu, že původní modely Altmana nebo Ohlsona jsou už přes 50 let staré, tak se část akademiků věnuje jejich aktualizaci vzhledem k novým podmínkám, přičemž další část se věnuje tvorbě zcela nových modelů.

Autoři, kteří tvoří nové modely, tak dělají přidáním nových vysvětlujících proměnných nebo pomocí změny formy modelu. Těmito úpravami se snaží zajistit vyšší predikční schopnost daného modelu oproti jeho existujícím alternativám.

V současnosti se tyto dva způsoby většinou kombinují ve tvaru tržních modelů (přidané tržní proměnné), které jsou založeny na principu oceňování opcí. K těmto autorům se řadí např. Shumway (2001), Hillegeist et al. (2004), Vassalou a Xing (2004) a další.

Druhý pomyslný směr, který se zabývá upravením původních modelů na současné podmínky, je stále veden původními autory, přičemž Ohlson svůj vlastní model upravil například v roce 2003 (Kubíčková 2015). Altman postupem času také pozměnil svůj původní model a v roce 1977 jej přepočítal pro společnosti, které nebyly součástí regulovaného kapitálového trhu v USA (Altman, 1977). Nedávno jej dokonce obohatil o makroekonomické determinanty a dal tak vzniknout tzv. Z-metrics (Altman et al., 2010). Nabízí se ale otázka, proč jsou vůbec jejich původní modely takto modifikovány.

Soudobí autoři se shodují na názoru, že účetní modely by se měly využívat pouze v období, které je blízké jejich vzniku. Jinými slovy, čím je testovaná společnost časově vzdálenější od původního vzorku společností, na kterém byl model vytvořen, tím více klesá predikční schopnost daného modelu. Platt a Platt (1990) ve své práci naznačili, že rozdíl v ekonomických podmínkách, které vznikly výše zmíněným časovým nesouladem, mohou vést ke změně:

- vztahu vysvětlované (bankrot) a vysvětlujících (poměrové ukazatele) proměnných
- průměrných hodnot vysvětlujících proměnných
- vztahu mezi vysvětlujícími proměnnými

Tyto změny přisuzovali změně v obchodním cyklu, korporátní strategii, konkurenci na trhu a technologii.

Podobně se také v dnešní době spekuluje o tom, jestli jsou původní účetní modely aplikovatelné na různá odvětví průmyslu a různé země. Příkladem autorů, kteří se touto problematikou zabývali, jsou např. Grice a Dugan (2003). Ti se ve své práci zaměřili jak na časový nesoulad, tak na použitelnost v jiných odvětvích průmyslu. Zjistili, že přepočítaný model dosahuje lepších výsledků než originál, čím potvrdili rozdílnost ekonomických podmínek vzhledem k času. Na jiném vzorku prokázali, že model není citlivý ke změně v odvětví průmyslu.

Tyto důvody jsou hlavní motivací pro úpravu původních účetních modelů. Úpravy většinou spočívají v převzetí původních tvarů proměnných a formy modelu (MDA, logit) a jejich aplikací na aktuální vzorek firem z dané země a období. Během této aplikace se často stává, že autoři některé z původních proměnných vyloučí z modelů díky statistické nevýznamnosti (většinou pomocí t-testu) a pro zbylé proměnné jsou vypočítány nové koeficienty. Takto upravené modely mají velmi dobré výsledky i na společnostech, které nebyly zahrnuty při tvorbě modelu (tzv. hold-out sample). Několik autorů společně s jejich modely bude zmíněno v podkapitole, která se zabývá modifikacemi těchto původních modelů.

## 2.2 Hlavní typy bankrotních modelů

V předchozí části byl rozebrán vývoj bankrotních modelů, přičemž byly zmíněny nejdůležitější typy těchto modelů, které dodnes ovlivňují většinu autorů v této oblasti. Tyto modely určitým způsobem přispěly k vývoji bankrotních modelů do dnešních podob, které se mnohdy od původních modelů liší, ale stále z nich vycházejí. Proto je účelné, aby v této práci byly zmíněny nejen obecné popisy těchto modelů, ale i jejich formy a okolnosti vzniku.

Historicky první model predikce bankrotu, který byl sestaven Fitzpatrickem v roce 1931, se v praxi neosvědčil, a tak nebude dále rozebírán. Jako první proto bude popsán první významný model z roku 1966, který byl vytvořen Beaverem. Beaver k sestavení bankrotního modelu využil 30 poměrových ukazatelů, které později omezil na finálních šest, s nejvýznamnějším vztahem k bankrotu společností. Tyto ukazatele jsou uvedeny v tabulce č. 1.

Tab. 1 Poměrové ukazatele použité Beaverem (1966)

<b>Beaver (1966)</b>	
X <sub>1</sub>	Peněžní tok / Cizí zdroje
X <sub>2</sub>	Čistý zisk / Aktiva celkem
X <sub>3</sub>	Cizí zdroje / Aktiva celkem
X <sub>4</sub>	Pracovní kapitál / Aktiva celkem
X <sub>5</sub>	Oběžná aktiva / Krátkodobé cizí zdroje
X <sub>6</sub>	(Defenzivní aktiva - krátkodobé cizí zdroje) / Provozní náklady bez odpisů

Vzhledem k faktu, že tento model není dnes využíván a ani pro účely této práce není klíčový, tak nebude dále specifikován. Naopak významným a v aktualizovaných formách dodnes používaným je Altmanův Z-score model z roku 1968.

Edward I. Altman se stejně jako jeho předchůdce Beaver zaměřil na poměrové ukazatele jako klíčové faktory v predikci bankrotu. Jeho přínos do této problematiky spočíval hlavně v přístupu k tvorbě formy modelu. Snažil se o vytvoření modelu, jehož výstupem by byla jediná hodnota, podle které by se dalo rozhodovat o finančním zdraví společnosti. Tohoto výsledku docílil použitím tzv. "MDA" metody (multiple discriminant analysis), což je vícenásobná diskriminační analýza. Jejím výstupem je lineární rovnice s vysvětlovanou proměnnou Z a několika vysvětlujícími proměnnými s přiřazenými koeficienty (váhami). Finální verze z roku 1968, vytvořená pro společnosti obchodované na kapitálových trzích, má následující formu.

$$Z = 1,2X_1 + 1,4X_2 + 3,3X_3 + 0,6X_4 + 0,999X_5 \quad (1)$$

Uvedené proměnné X<sub>1</sub> až X<sub>5</sub> jsou zastoupeny vybranými poměrovými ukazateli, které jsou z většiny získány z účetních výkazů společnosti. Výjimkou je proměnná X<sub>4</sub>, která využívá tržní hodnotu vlastního kapitálu. Přesná forma těchto proměnných je popsána v tabulce č. 2.

Tab. 2 Poměrové ukazatele použité Altmanem (1968)

<b>Altman (1968)</b>	
X <sub>1</sub>	Čistý pracovní kapitál / Aktiva celkem
X <sub>2</sub>	Nerozdělený zisk / Aktiva celkem
X <sub>3</sub>	EBIT / Aktiva celkem
X <sub>4</sub>	Tržní hodnota vlastního kapitálu / Závazky celkem
X <sub>5</sub>	Tržby / Aktiva celkem

Jejich přiřazené koeficienty jsou vypočítány modelem a měly by v této formě zastupovat nejvhodnější váhu jednotlivých ukazatelů při posuzování rizika bankrotu společnosti. Hodnota  $Z$  v tomto modelu zastupuje vysvětlovanou neznámou a je nutné ji přiřadit intervaly, které by ji interpretovaly uživatelům. Tyto zóny jsou Altmanem slovně popsány jako bezpečná (safe) zóna, zóna nouze (distress) a šedá (grey) zóna. Společnosti, jejichž  $Z$  hodnota spadá do bezpečné zóny, čelí minimálnímu riziku, kdežto společnosti v zóně nouze čelí reálnému riziku bankrotu. Šedá zóna představuje nejasné pásmo, ve kterém se nedá o společnosti jednoznačně tvrdit, jestli je bankrotem ohrožena nebo ne. Jednotlivé intervaly (zóny) pro  $Z$  hodnotu jsou vypsány níže (Salimi, 2015):

Tab. 3 Zóny Altmanova Z-score (1968)

Zóny pro Z hodnotu		
$Z >$	2,99	Bezpečná zóna
$Z \in$	$<1,8; 2,99>$	Šedá zóna
$Z <$	1,8	Zóna nouze

Altmanův model v této podobě z roku 1968 čelí v současnosti kritice. Kritika se hlavně zaměřuje na použitelnost modelu na datech, která nejsou z období, ve kterém byl model vytvořen. Jinými slovy je zpochybnitelné, jestli model ve své původní podobě z roku 1968 bude stejně dobře předpovídat bankrot i u dat ze současné doby. Příkladem takového kritika je např. Grice (2001), který nejen kritizuje použitelnost v souvislosti s časem, ale také použitelnost modelu v jiných průmyslových odvětvích. Podle něj se ekonomické podmínky jako inflace, úroková míra a dostupnost úvěru časem mění a mohou negativně ovlivnit predikční schopnost původního Altmanova modelu.

Dalším významným modelem, který je dodnes hojně využíván, je model J. A. Ohlsona z roku 1980. Ohlson sám naznačil, že se nesnažil o vývoj nového modelu, co se proměnných týče, ale pouze použil proměnné (poměrové ukazatele), které se nejčastěji objevovaly v literatuře. Dodnes se věří, že jako první k predikci bankrotu použil logistickou regresi (občas označovanou jako logit model). Jeho model odráží situaci prostředí v USA během 80. let minulého století. Data získal právě z amerických firem v letech 1970–1976. Testovaný vzorek společností obsahoval 105 zbankrotovaných a 2058 zdravých firem. Na základě těchto dat sestavil rovnici o devíti vysvětlujících proměnných v následující podobě:

$$Q = -1,3 - 0,4X_1 + 6,02X_2 - 1,4X_3 + 0,1 X_4 - 2,4X_5 - 1,8X_6 + 0,3X_7 - 1,7X_8 - 0,5X_9 \quad (2)$$

Vysvětlující proměnné v tomto modelu mají charakter poměrových ukazatelů a jejich přesný tvar je uveden v tabulce č. 4.

Tab. 4 Poměrové ukazatele použité Ohlsonem (1980)

<b>Ohlson (1980)</b>	
X <sub>1</sub>	log (Aktiva celkem / deflátor HDP)
X <sub>2</sub>	Závazky celkem / Aktiva celkem
X <sub>3</sub>	Čistý pracovní kapitál / Aktiva celkem
X <sub>4</sub>	Krátkodobé cizí zdroje / Oběžná aktiva
X <sub>5</sub>	$X_5 = \begin{cases} 1, & \text{Závazky celkem} > \text{Aktiva celkem} \\ 0, & \text{Závazky celkem} \leq \text{Aktiva celkem} \end{cases}$
X <sub>6</sub>	Čistý zisk / Aktiva celkem
X <sub>7</sub>	(Čistý zisk + odpisy) / Závazky celkem
X <sub>8</sub>	$X_8 = \begin{cases} 1, & (NI_t + NI_{t-1}) < 0 \\ 0, & (NI_t + NI_{t-1}) \geq 0 \end{cases}$
X <sub>9</sub>	$(NI_t + NI_{t-1}) / ( NI_t  +  NI_{t-1} )$

*Pozn.  $NI_t$  je čistý zisk aktuálního období a  $NI_{t-1}$  je čistý zisk minulého období*

Výsledná hodnota Q je pouze mezivýsledek, který se musí dosadit do rovnice pro výpočet pravděpodobnosti:

$$P = \left( \frac{1}{1 + e^{-Q}} \right) \quad (3)$$

Výsledkem po dosazení je pravděpodobnost bankrotu dané firmy v předem určeném časovém intervalu (jeden rok předem, dva roky, ...). Tato hodnota se pohybuje v intervalu  $<0,1>$ . Z tohoto vztahu také vyplývá, že čím je hodnota Q vyšší, tím je vyšší pravděpodobnost bankrotu, kdežto pokud je hodnota Q nízká, tak situace v podniku je poměrně stabilní. Tato situace se dá zapsat vztahem:

$$\begin{aligned} \text{Pokud } Q < 0, & \text{ tak } P \rightarrow 0 \\ \text{Pokud } Q > 0, & \text{ tak } P \rightarrow 1 \\ \text{Pokud } Q = 0, & \text{ tak } P = 0,5 \end{aligned}$$

Značnou výhodou logit modelu (model založený na logistické regresi) je výsledek uvedený přímo v pravděpodobnosti bankrotu dané firmy. Ohlsonem určená hodnota 0,5 zde značí pomyslnou linii mezi zdravou společností a společností ve finanční tísní. Interval  $<0,45; 0,55>$  definuje šedou zónu, která zajistí, aby firmy na pomezí hodnoty 0,5 nebyly mylně klasifikovány do špatné kategorie.

Poslední významný model, z pomyslného směru účetních modelů, který bude okrajově zmíněn, je model Marka E. Zmijewskiho z roku 1984. Stejně jako Ohlson,



se Zmijewski nepokoušel začlenit nové proměnné, ale použil ty, které se nejčastěji vyskytovaly v literatuře. Další podobností s Ohlsonem je i změna formy modelu, kdy Zmijewski využívá k predikci bankrotu probit model. Jeho model je formulován na vzorku 40 zbankrotovaných a 800 zdravých společností z let 1972–1978. Finální model, který na základě těchto dat vznikl, má následující podobu:

$$X = -4,3 - 4,5X_1 + 5,7X_1 - 0,004X_3 \quad (4)$$

Použité proměnné a jejich podoba jsou uvedeny v tabulce č. 5.

Tab. 5 Poměrové ukazatele použité Zmijewskim (1984)

<b>Zmijewski (1984)</b>	
X <sub>1</sub>	Čistý zisk / Aktiva celkem
X <sub>2</sub>	Závazky celkem / Aktiva celkem
X <sub>3</sub>	Oběžná aktiva / Krátkodobé cizí zdroje

V současné době se stále populárnějšími stávají i modely založené na teorii oceňování opcí. Nejčastěji využívaným modelem v tomto směru je pravděpodobně Black-Scholesův model s různými modifikacemi. I když tato forma modelu nebude využita pro zpracování data setu v této práci, tak je v současnosti dostatečně významná, aby byly popsány její principy a tvorba.

Jednou z hlavních myšlenek tohoto modelu, kterou ve své práci vypořádal Merton (1974), je posouzení vlastního kapitálu společnosti jako call opce na firemní aktiva. Pokud společnost zbankrotuje, tak držitelé vlastního kapitálu nárokují pozůstatky aktiv a jsou předmětem pouze omezeného ručení, čímž imitují vyplacení opcí. S přihlédnutím k formě BSM je realizační cena call opce rovna nominální hodnotě firemních závazků a opce vyprší v čase T, kdy dozraje dluh. Držitelé vlastního kapitálu v čase T uplatní opci a pokud je hodnota firemních aktiv vyšší než nominální hodnota jejich závazků, tak vyplatí věřitele. V opačném případě, kdy hodnota aktiv nestačí na pokrytí dluhu společnosti, držitelé vlastního kapitálu nechají opci vypršet. V takovém případě firma zažádá o bankrot, její vlastnictví je převedeno na věřitele a vyplacení držitelů vlastního kapitálu je nulové. Pravděpodobnost obou variant udává cenu dané call opce a je zahrnuta v samotném BSM modelu.

Hillegeist et al. (2004) ve své práci uvádí rovnici pro ocenění vlastního kapitálu jako evropské opce:

$$V^E = V_A e^{-\delta T} N(d_1) - X e^{-rT} N(d_2) + (1 - e^{-\delta T}) V_A \quad (5)$$

kde,  $N(d_1)$  a  $N(d_2)$  jsou standardní normální rozdělení  $d_1$  a  $d_2$ , které mají tvar:

$$d_1 = \frac{\ln\left(\frac{V_A}{X}\right) + \left(r - \delta + \frac{\sigma_A^2}{2}\right)T}{\sigma_A \sqrt{T}} \quad (6)$$

a

$$d_2 = d_1 - \sigma_A \sqrt{t} = \frac{\ln\left(\frac{V_A}{X}\right) + \left(r - \delta + \frac{\sigma_A^2}{2}\right)T}{\sigma_A \sqrt{T}} \quad (7)$$

V tomto modelu  $V_E$  představuje hodnotu vlastního kapitálu,  $V_A$  současnou tržní hodnotu aktiv,  $X$  nominální hodnotu dluhu dozrávajícího v čase  $T$ ,  $r$  bezrizikovou úrokovou míru,  $\delta$  míru dividend vyjádřených pro  $V_A$  a  $\sigma$  je směrodatná odchylka rentability aktiv. V rámci BSM modelu je pravděpodobnost bankrotu jednoduše pravděpodobnost toho, že současná hodnota aktiv ( $V_A$ ) je nižší než nominální hodnota závazků ( $X$ ) v čase  $T$  (tzn.  $V_A(T) < X$ ).

Finální rovnice pro výpočet pravděpodobnosti  $V_A(T) < X$  byla představena v práci McDonalda (2002) a má tento tvar:

$$N\left(\frac{\ln\frac{V_A}{X} + \left(\mu - \delta - \frac{\sigma_A^2}{2}\right)T}{\sigma_A \sqrt{T}}\right) = \text{BSM - Prob} \quad (8)$$

Tato rovnice ukazuje, že pravděpodobnost bankrotu je funkce, která řeší poměr současné hodnoty firemních aktiv a nominální hodnoty jejich závazků ( $V_A/X$ ), upravený pro očekávaný nárůst v hodnotách aktiv ( $\mu - \delta - \sigma_A^2/2$ ), vzhledem k volatilitě aktiv ( $\sigma_A$ ).

Odborná literatura se v současnosti zabývá i novými způsoby predikce bankrotu, jako je např. využití neuronových sítí, což je v podstatě využití umělé inteligence. Ale vzhledem k tomu, že tyto pokusy jsou i dnes zmiňovány velmi zřídka a často jen okrajově, tak v této práci nebudou rozebrány.

### 2.3 Modifikace O-score a Z-score ve světové literatuře

V předchozí podkapitole byl zmíněn původní Ohlsonův model a jeho tvar z roku 1980. Současná literatura se ale víceméně shoduje na názoru, že tento původní model je v současné době těžko použitelný a minimálně vyžaduje modifikaci koeficientů, aby měl dostatečnou predikční schopnost. Proto se tato podkapitola zaměří

na aktualizované verze jak Ohlsonova, tak i Altmanova modelu, a to jak tvůrci samotnými, tak dalšími autory, kteří tyto modely modifikovali.

V roce 1993 byl otestován Ohlsonův i Altmanův původní model na soudobých společnostech, aby se zjistilo, jestli se jejich parametry změní vzhledem k původním variantám (Jin, 1993). Byla tedy vypracována jeho nová verze, získaná z dat 99 bankrotujících a 1980 prosperujících společností během let 1981-1990. Oblasti financí a přepravy byly cíleně z tohoto vzorku vynechány. Nakonec byly zkonstruovány dvě varianty, a to varianta predikující bankrot jeden rok dopředu a druhá dva roky předem. V tomto modelu byly využity všechny výpočty i proměnné z původního modelu a změny nastaly pouze v koeficientech a konstantě. V dalších letech se potvrdilo, že predikční schopnost modelu klesá s větším časovým odstupem a tento model přestal být vhodným pro predikci.

Díky tomu se jej rozhodl Ohlson znovu přepočítat. Při tvorbě nového modelu se rozhodl pro mírnější variantu, než je predikce bankrotu a predikoval finanční tíseň společnosti, která byla dle jeho názoru užitečnější. Nový model byl tedy derivován v roce 2003 na relativně velkém vzorku společností z prostředí USA. Tento model měl tři varianty, a proto byly vytvořeny tři nové sady koeficientů původních devíti proměnných. První varianta (A) byla vytvořena z celého vzorku společností, druhá (B) pouze ze vzorku společností s finančními problémy a třetí (C) ze společností patřících do prostředí průmyslu. Oproti modelu z roku 1993, byla vynechána konstanta a po testování se zjistilo, že tento model předčil veškeré dřívější varianty. Znovu se tedy potvrdilo, že modely přepočítané na současné podmínky mají lepší predikční schopnosti než jejich starší verze (Wang et al., 2010). Tento model se dá považovat za poslední mezinárodní variantu, protože jeho další verze byly přepočítány vzhledem ke specifickému prostředí jednotlivých zemí.

Jednou z takových prací je studie Muzira a Çağlara (2009) z tureckého prostředí. Autoři zde vycházeli z poměrně malého vzorku společností (70). Opět se nezměnily samotné proměnné, ale pouze dospěli k jiným hodnotám koeficientů.

V roce 2010 se o přepočítání vah jednotlivých proměnných pokusili i australští ekonomové z univerzity v Queensland (Wu, Gaunt a Gray; 2010). Jejich práce byla originální díky robustnosti jejich data setu, který získali z 50 611 společností. Použili data z let 1980 až 2006, přičemž 49 724 společností bylo zdravých a 887 bylo bankrotujících. Výpočtem získali nový model, který obsahoval všechny původní proměnné. S předchozí tureckou verzí se opět lišil pouze v hodnotách koeficientů (znaménka jednotlivých proměnných byla identická).

V témže roce se americký profesor Michael Campbell spojil s čínským ekonomem Ying Wangem a spoluprací se snažili vytvořit nový model pro čínskou ekonomiku. Jejich data set derivovali z čínských společností v letech 1998 až 2008 a stejně jako u původního Ohlsonova modelu dospěli ke třem variantám modelu s různými dobami predikce bankrotu. První varianta predikuje na rok dopředu, druhá na dva a třetí buď na jeden, nebo dva roky předem. Vzhledem k tomu, že použili všechny původní proměnné, tak pochybovali o tom, jestli je tento přepočítaný model nejlepší variantou. Proto pro jistotu vytvořili novou sadu dalších tří přepočítaných variant modelu, ale v nich už nepoužili všechny původní proměnné.

V tomto novém modelu chtěli zvýšit jeho predikční schopnost a aplikovatelnost tím, že omezili proměnné na pět nejdůležitějších. (Wang et al., 2010).

V roce 2011 se v Íránu snažili otestovat použitelnost čtyř nejznámějších predikčních modelů, kterými jsou Altman, Ohlson, Shumway a Zmijewski. Během jejich přepočítání tak vytvořili i novou verzi Ohlsonova modelu pro danou zemi. V použitém data setu bylo 1500 společností, ze kterých 142 bylo ve finanční tísní. Nový model obsahoval všechny původní proměnné i konstantu a znovu tedy byla přepočítána pouze váha jednotlivých proměnných (Kordlar a Nikbakht, 2011). Veškeré výše zmíněné modifikace Ohlsonova modelu jsou uvedeny v souhrnné tabulce v přílohách, v sekci Literární přehled.

Druhým typem modelu, který bude využit pro zpracování dat v této práci je Altmanův Z-score model a jeho modifikace, proto je důležité uvést alespoň nejdůležitější proměny tohoto modelu.

Po roce 1968 bylo zjištěno, že původní Altmanův model není časově flexibilní, a tak začal být ostatními autory upravován. Vzhledem k těmto úpravám a kritikám jeho modelu, začal sám Altman revidovat a vytvářet nové modely, které by nejlépe vyhovovaly měnícím se podmínkám na trhu. V této aktivitě pokračuje dodnes a ve spolupráci s dalšími autory vytvořil několik průlomových modifikací, které zde budou popsány.

Za první důležitou modifikaci je považován ZETA model z roku 1977. ZETA model byl vytvořen jako reakce na neustálou poptávku po bankrotním modelu, který by byl vhodný pro soukromé společnosti, jež nejsou obchodovány na kapitálovém trhu. Obecně vzato je to tedy původní model z roku 1968 modifikovaný pro tento typ společností. Daná modifikace spočívala hlavně v proměnné  $X_4$ , kde byla tržní hodnota vlastního kapitálu nahrazena účetní hodnotou vlastního kapitálu. Altmanův ZETA model má tedy následující tvar (Altman, 1977):

$$Z' = 0,717X_1 + 0,847X_2 + 3,107X_3 + 0,420X_4 + 0,998X_5. \quad (9)$$

Se změnou modelu musel Altman poupravit i dosavadní kritéria (intervaly) pro hodnocení výsledné hodnoty. Nové intervaly posuzují společnost s hodnotou nad 2,9 jako zdravou, a naopak společnost s hodnotou pod 1,23 posoudí jako bankrotující. Hodnoty v intervalu (1,23; 2,9) jsou tzv. šedé pásmo, kde nelze jasně posoudit finanční zdraví podniku. Je zřetelné, že šedé pásmo pro ZETA model je mnohem širší než pro původní model (1,8; 2,9) z roku 1968.

V roce 1995 byl vytvořen další model, který se zabýval nevýrobními firmami a rozvíjejícími se trhy (emerging markets). V této modifikaci se neuvážovalo s proměnnou  $X_5$ , která určuje produkční sílu a má tvar (tržby celkem/aktiva celkem). Modifikovaný model vypadá následovně:

$$Z'' = 6,56X_1 + 3,26X_2 + 6,72X_3 + 1,05X_4 \quad (10)$$

Pro nový model musely vzniknout i nové vysvětlující intervaly:

Z >	2,6	Malá pravděpodobnost bankrotu
Z ∈	<1,1; 2,6>	Šedá zóna
Z <	1,1	Vysoká pravděpodobnost bankrotu

Opět lze pozorovat rozšíření šedé zóny jak oproti původnímu modelu, tak oproti ZETA modelu.

Další důležitou modifikací samotného Altmana je model z roku 2010 nazvaný Z-metrics, který vznikl ve spolupráci se skupinou RiskMetrics. Při tvorbě daného modelu Altman et al. přistoupili k několika zásadním změnám, které dosud v předchozích modelech Z-score nebyly použity. První důležitou změnou je změna formy modelu, kdy Altman využil logistickou regresi a Z-metrics má tedy tvar logit modelu. Další změna se týká využitých proměnných, kde je využito nejen účetních, ale i tržních (hodnota akcií, úrokový spread atd.) a makroekonomických (růst HDP, nezaměstnanost atd.) proměnných, které mají přispět k lepší predikci. Poslední a velmi důležitou změnou je výstup modelu. Nejedná se totiž o klasický bankrotní model, ale o model ratingový, a tudíž jeho výstupem je rating společnosti. Rating je rozdělen na tři stupně (nízký, střední, vysoký) a nabývá hodnot od ZA+ po ZF-. Kritériem pro zařazení společnosti do dané kategorie ratingu je pravděpodobnost úpadku. Tento model je vytvořen pro velké společnosti obchodované na kapitálových trzích, velké neobchodované společnosti, malé obchodované společnosti v USA a Kanadě, ale také pro velké a malé společnosti mimo Kanadu a USA.

Příklad ratingové tabulky výše zmíněného Z-metrics uvedla ve své práci i Dluhošová (2011) a je vyobrazen v přílohách (část Literární přehled).

## 2.4 Modifikace vybraných modelů v české literatuře

V předchozí podkapitole byly popsány aktuální modifikace Z-score a O-score modelu, které poslouží pro další výpočty v této práci. Tyto modifikace pocházeli od světových autorů, ale to neznamená, že v české literatuře se toto téma nevyskytuje. Stejně jako např. v Turecku, Číně či Íránu, tak i zde se mnoho autorů snažilo otestovat a přizpůsobit tyto dva významné modely pro domácí prostředí.

Ze současných prací je možno zmínit Adamce (2010), který pracoval na porovnání několika predikčních modelů – IN05, ZETA, Ohlsonova a Shumwayova modelu. Z výsledků ale zjistil, že rok před bankrotem byly všechny znaky bankrotu tak zřetelné, že všechny zmíněné modely dokázaly predikovat bankrot s vysokou přesností.

Ve stejném roce byla publikována studie Klečky a Scholleové (2010), ve které se zaměřili na sklářský průmysl a snažili se jej zhodnotit pomocí Z-score i IN05. Z výsledků své práce vyvodili, že ani jeden z modelů nedokázal predikovat bankrot dostatečně dopředu, ale s určitou přesností dokázali predikovat finanční tíseň jednotlivých společností. V roce 2012 na tuto práci navázala Čámská a Hájek (2012), kteří znovu otestovali finanční zdraví společností z celého sklářského odvětví, ovšem pouze za použití Altmanova a IN05 modelu.

Šlégr (2013) porovnal výsledky Z-score a IN05 modelů na vzorku dat z padesáti největších firem České Republiky v letech 2006 až 2010. Ukázal, že Z-score

model má lepší predikční schopnost než IN05, i když soudobá světová literatura takové výsledky nepotvrdila.

Kupilík (2013) testoval původní i modifikované verze Ohlsonova modelu a potvrdil, že novější verze mají lepší predikční schopnosti. Toto tvrzení je často zmiňováno i ve světové literatuře a je jedním z hlavních důvodů, proč se akademici zabývají úpravami těchto modelů.

Jedna z nejnovějších prací z této oblasti v domácím prostředí je studie Kubíčkové (2015), která se zaměřila na predikční schopnosti Ohlsonova logit modelu (O-score) na datech 1996 malých a středních podniků v ČR. Pro porovnání přidala i IN05<sup>1</sup>, Z-score a Tafflerův model.<sup>2</sup> Na základě výsledků této práce ukázala, že Ohlsonův a Tafflerův model mají velmi podobné výsledky, přičemž asi 90 % firem bylo klasifikováno jako zdravé. Na druhou stranu Altmanovo Z-score mělo velmi podobné výsledky s modelem IN05, kdy oba dva modely označily cca 40 % firem jako zdravé. Ze stejných výsledků lze vyčíst, že Z-score i IN05 mají při porovnání o dost více firem klasifikovaných v šedé zóně.

Valecký a Slivková (2012) se zaměřili na oblast predikce bankrotu z jiného úhlu pohledu. Místo rekalkulace starších známých modelů, se pokusili o vytvoření vlastního mikroekonomického modelu úpadku ve formě logit modelu. To znamená, že i tvar jejich proměnných se liší od proměnných běžně využívaných v modelech o-score a z-score. Např. často využívaný ukazatel celkové zadluženosti zde byl změněn na zadluženost běžnou a stejně tak i ukazatel běžné likvidity byl upozaděn likviditou okamžitou. Jejich model byl vytvořen na datech 400 společností z roku 2008. Samotní autoři zdůrazňují, že model byl vytvořen na krátkém časovém intervalu, a proto by neměl být využit pro vzdálenější budoucnost.

---

<sup>1</sup> Index IN05 (Neumaierová a Neumaier, 2005) je kombinací bankrotního a bonitního modelu, který byl vytvořen manželi Neumaierovými v roce 2005. Tento model byl vytvořen pro prostředí ČR.

<sup>2</sup> Tafflerův model (Taffler, 1983), je bankrotní model z roku 1983 vytvořen pro prostředí Spojeného království.

## 3 Výběr dat a proměnných

Tato podkapitola se zabývá výběrem vzorku společností (tzv. data setem), na kterém budou dále konstruovány predikční modely bankrotu. Tento data set se skládá ze společností, které jsou uvedeny v databázi Amadeus od společnosti Bureau van Dijk. Přesněji se v práci využijí účetní výkazy těchto společností, ze kterých lze vypočítat žádané poměrové ukazatele.

### 3.1 Kritéria pro výběr společností

Vzhledem k tomu, že databáze Amadeus (stejně jako většina podobných databází) shromažďuje značné množství rozdílných společností, tak je nutné určit kritéria výběru pro specifický vzorek, s kterým bude pracováno. Toto omezení je zavedeno vzhledem k domněnce, že na užším přesně specifikovaném vzorku budou lépe pozorovatelné trendy, které ovlivňují finanční situaci těchto společností, a tím umožní lepší predikci bankrotu. Finální výběr kritérií pro výběr data setu má následující podobu:

- Země původu
- Hospodářské odvětví
- Velikost společnosti
- Právní forma společnosti
- Stav společnosti

První kritérium není třeba rozvádět a ze své podstaty je jasné, že se jedná o zemi původu dané společnosti. I když se toto kritérium zdá být jasné a možná i zbytečné, tak je nutné jej aplikovat, a to kvůli rozsáhlosti dané databáze. Databáze Amadeus totiž shromažďuje nejen údaje českých firem, ale téměř všech evropských zemí.

Druhé kritérium je už mnohem konkrétnější a je aplikováno v návaznosti na zjištění několika předchozích prací (např. Grice a Dugan, 2003), ve kterých se zjistilo, že model vytvořený pro specifické odvětví má pro toto odvětví lepší predikční schopnost než obecně zaměřené modely. Proto bude v této práci zahrnuto pouze jedno odvětví, kterým je stavebnictví. Stavebnictví bylo vybráno, protože se u něj předpokládá vyšší náchylnost k bankrotu. Tato domněnka se zakládá na faktu, že vývoj celého stavebního odvětví je náchylný na změny hospodářského cyklu a mnohdy i na změny sezónní. Proto se očekává, že středně velké společnosti často nedisponují dostatečnými prostředky pro překlenutí těchto útlumů a jsou více náchylné k vyhlášení bankrotu.

Další kritérium je rozčlenění společností v daném oboru do tří obecných skupin – velké, střední a malé společnosti. Toto kritérium je aplikováno z jednoho důvodu, a to obecné použitelnosti výsledného modelu. Pro odůvodnění tohoto kritéria je nutné představit si situaci velké a malé stavební firmy. Velké firmy, které podnikají s ročními obraty v miliardách a s tisíci zaměstnanci mají jiné podmínky než zbytek daného odvětví. Jejich vyjednávací pozice s dodavateli i odběrateli je na

jiné úrovni a mnoho situací, které by mohly být pro malé a střední firmy rizikové je možné u těchto firem s lehkostí vyřešit už jen díky reputaci dané společnosti. Dalším důvodem je vztah s bankami – je totiž jasné, že banky budou k takové společnosti vlídnější, ať už se bude jednat o výši úroků nebo podmínkách placení. U predikce finanční tísně nebo bankrotu těchto firem je lepší řešit každou firmu individuálně, a proto nebudou do této práce zahrnuty. Malé firmy na druhou stranu zažívají opačné situace. Mají malou vyjednávací pozici u dodavatelů i bank, jejich provoz je dražší (malé úspory z rozsahu, vyšší úroky u úvěrů, ...) a některé situace, které by byly pro střední a velké firmy zanedbatelné, mohou znamenat pro malou firmu existenční krizi. Další problém těchto společností tkví v managementu. Vzhledem k malým finančním možnostem je málokdy zaměstnán kvalifikovaný manažer a mnohdy celý management zastává sám vlastník firmy. V případě, že neoplývá znalostmi potřebnými pro řízení provozu a plánování do budoucna, tak bankrot může nastat i díky selhání lidského faktoru, a proto i tyto firmy budou z data setu vyřazeny. V potaz jsou tedy brány pouze středně velké společnosti, které mají na základě evropských norem tento charakter: počet zaměstnanců v intervalu <50; 250) a roční obrát v intervalu <10; 50) milionů eur.

Čtvrté výběrové kritérium je zaměřeno na právní formu společností, přičemž jeho smyslem je eliminace akciových společností. Tyto společnosti je třeba vyřadit už jen kvůli možnosti jejich zániku, který nemusí mít nic společného s úspěšností společnosti jako takové. Příkladem může být „nepřátelské převzetí“, kdy v rámci konkurenčního boje skoupí konkurenční firma většinový podíl dané společnosti a buď ji nechá zaniknout nebo začlenit do svého majetku pod jiným jménem.

Poslední kritérium je stav společnosti, který vypovídá o tom, jestli je firma aktivní nebo jestli zanikla. Veškeré společnosti tedy roztrídí do dvou souborů, a to do souboru „zdravých“ (nezaniklých) firem a do souboru bankrotních firem. Komplikace nastává ve chvíli, kdy je potřeba určit co přesně znamená firma, která zažívá finanční tíseň. Ve světové literatuře se často používá tento popis (Grice & Dugan, 2003):

- Společnost, která byla vyřazena z dané databáze společností kvůli bankrotu
- Společnost, která byla vyřazena z dané databáze společností kvůli likvidaci
- Společnost, kterou agentura S&P ohodnotila ratingem CCC a níže
- Společnost, jejíž akcie dostaly rating „nižší B“ a níže

Vzhledem k tomu, že v této práci je brán v úvahu bankrot společností, jenž má užší význam než jen finanční tíseň, tak ratingové charakteristiky v hodnocení společností budou opominuty. Databáze Amadeus nabízí několik možností, jak rozlišit bankrotní firmy, ale pro účely této práce byly vybrány tyto typy:

- Společnost, která vyhlásila bankrot
- Společnost v likvidaci
- Společnost, která vyhlásila bankrot a zanikla
- Společnost v likvidaci, která zanikla



I když se tyto typy bankrotů mohou zdát velmi podobné, tak je v nich zásadní rozdíl. Ten spočívá v tom, jestli daná společnost využila vyhlášení bankrotu nebo likvidaci ke splacení svých závazků, a poté pokračovala ve své činnosti (i když v omezené míře) nebo jestli tímto aktem zanikla.

### 3.2 Výběr vysvětlujících proměnných

Po určení kritérií pro výběr žádaných společností, je nutné vybrat i žádané vysvětlující proměnné. Jak už bylo v práci zmíněno dříve, tak tyto proměnné se v zásadě mohou objevit ve třech podobách – účetní, tržní nebo makroekonomické (někdy i specifické) proměnné. Při výběru společností byl zavržen typ akciové společnosti, a proto není nutné uvažovat o proměnných tržních a v úvahu připadají pouze proměnné účetní a makroekonomické. Vzhledem k tomu, že tato práce se silně inspirovuje dvojicí autorů Ohlson a Altman, tak je vybrán stejný typ proměnných, které byly zahrnuty do jejich modelů, což jsou proměnné účetní.

Účetní proměnné získaly svůj název ze způsobu, jakým jsou získány. V drtivé většině se tyto proměnné vyskytují v podobě poměrových ukazatelů a je proto potřeba využít účetních výkazů dané společnosti, jako je rozvaha nebo výkaz zisku a ztrát, pro jejich výpočet.

Jelikož je nutné nezanedbat žádnou potenciálně významnou proměnnou, tak budou do množiny proměnných zahrnuty veškeré proměnné z původního Ohlsonova a Altmanova modelu. Původní Ohlsonovy proměnné z roku 1980 jsou vypsány v tabulce č. 3 a Altmanovy z roku 1968 v tabulce č. 2.

V podkapitole o modifikacích těchto modelů byly zmíněny pozdější modifikace, které většinou spočívaly ve vypuštění některé z proměnných z modelu, díky její statistické nevýznamnosti. Ale vzhledem k faktu, že i tyto modifikace jsou poměrně zastaralé, tak je možné, že tyto vypuštěné proměnné mohou být znovu významné, a proto je nutné do testovaného vzorku zařadit všechny původní proměnné.

Jedinou výjimkou je změna proměnné  $X_4$  z původního Altmanova modelu (1968), kde je brána v úvahu tržní hodnota vlastního kapitálu. Vzhledem k absenci akciových společností, bude tato tržní hodnota změněna na účetní hodnotu (stejně jako v modelu Zeta z rovnice č. 9).

Mnoho ukazatelů z uvedených modelů jsou si smyslem podobné, a proto se předpokládá, že i jejich vliv bude podobný, ale vzhledem k tomu, že důkazy k těmto předpokladům budou nalezeny až samotným otestováním dat modelem, tak tento problém bude řešen v praktické části této práce.

Jediný ukazatel, který je totožný pro oba dva modely, je ukazatel čistý pracovní kapitál/aktiva celkem. Tento ukazatel, který staví čistý pracovní kapitál oproti celkovým aktivům, bude proto logicky použit jen jednou.

Poslední změnou je vypuštění druhého ukazatele z původního Altmanova modelu. Tento ukazatel vztahuje nerozdělený zisk vůči celkovým aktivům a představuje specifický tvar ukazatelů rentability z klasické finanční analýzy. Důvod pro vypuštění tohoto ukazatele z modelu je čistě praktický – databáze Amadeus neumožňuje tento údaj získat (přímo ani výpočtem z ostatních položek vlastního kapitálu). S přihlédnutím k faktu, že do modelu jsou zahrnuty i klasické tvary rentabili-

ty aktiv, které používají EBIT a čistý zisk, tak je pravděpodobné, že vypuštění této proměnné bude mít na celkový model zanedbatelný vliv.

V následující tabulce je uveden finální výběr všech proměnných, které budou zahrnuty do testování modelu v této práci.

Tab. 6 Vybrané vysvětlující účetní proměnné

<b>Vysvětlující proměnné</b>	
X <sub>1</sub>	Čistý pracovní kapitál / Aktiva celkem
X <sub>2</sub>	EBIT / Aktiva celkem
X <sub>3</sub>	Účetní hodnota vlastního kapitálu / Závazky celkem
X <sub>4</sub>	Tržby / Aktiva celkem
X <sub>5</sub>	log (Aktiva celkem / deflátor HDP)
X <sub>6</sub>	Závazky celkem / Aktiva celkem
X <sub>7</sub>	Krátkodobé cizí zdroje / Oběžná aktiva
X <sub>8</sub>	$X_8 = \begin{cases} 1, & \text{Závazky celkem} > \text{Aktiva celkem} \\ 0, & \text{Závazky celkem} \leq \text{Aktiva celkem} \end{cases}$
X <sub>9</sub>	Čistý zisk / Aktiva celkem
X <sub>10</sub>	(Čistý zisk + odpisy) / Závazky celkem
X <sub>11</sub>	$X_{11} = \begin{cases} 1, & (NI_t + NI_{t-1}) < 0 \\ 0, & (NI_t + NI_{t-1}) \geq 0 \end{cases}$
X <sub>12</sub>	$(NI_t + NI_{t-1}) / ( NI_t  +  NI_{t-1} )$

Pozn.: Označení  $NI_t$  je ekvivalentem pro čistý zisk v posledním období.  $NI_{t-1}$  potom logicky popisuje čistý zisk z období předešlého.

### 3.2.1 Odhad vlivu vysvětlujících proměnných

Po výběru kritérií data setu a tvaru vysvětlujících proměnných, které na něj budou aplikovány, je vhodné, aby byly tyto proměnné (poměrové ukazatele) obecně popsány a byl odhadnut jejich očekávaný vliv na finální výsledek modelu. Jinými slovy se jedná o určení toho, v jaké situaci (hodnotě) má daná proměnná na výsledek modelu kladný nebo záporný vliv (znaménko příslušného koeficientu ve výsledné rovnici).

#### Čistý pracovní kapitál/Aktiva celkem ... (X<sub>1</sub>)

První proměnná vztahuje čistý pracovní kapitál (pracovní kapitál očištěn o krátkodobé závazky) k celkovým aktivům a je ukazatelem finanční analýzy, jenž spadá do méně známé kategorie analýzy pracovního kapitálu. V zásadě vypovídá o tom, jaké je procentuální zastoupení čistého pracovního kapitálu na celkových aktivech. Podle Scholleové (2008) je např. u výrobních firem vhodné, aby tento ukazatel dosahoval hodnot mezi 10–15 %. Tato hodnota totiž vypovídá o tom, že značná část prostředku ve společnosti je vázána v dlouhodobých aktivech, jako jsou budovy nebo stroje. Vzhledem k nízkým doporučeným hodnotám ukazatele se odhaduje negativní efekt, a proto by jeho znaménko v rovnici mělo být kladné.

**EBIT/Aktiva celkem ... (X<sub>2</sub>)**

Druhá proměnná je klasickou formou jednoho z typů ukazatelů rentability – přesněji rentability aktiv (ROA). Klasické ukazatele rentability aktiv operují s čistým ziskem (9. proměnná) nebo se ziskem před zdaněním a nákladovými úroky, tedy EBIT a liší se tím, pro koho jsou určeny. Při použití EBIT jako v této proměnné, je ukazatel ROA v podstatě určen pro věřitele a stát, kdežto při užití čistého zisku má nejvyšší vypovídací hodnotu pro samotné vlastníky společnosti.

Nicméně obě dvě výše uvedené formy mají téměř stejný vliv na finanční situaci firmy, a to pozitivní. Jinými slovy se dá říci, že čím větší je ROA v jakékoliv formě, tím je to pro společnost výhodnější. Odhadované znaménko obou ukazatelů je tedy záporné. Proměnná č. 9 už tedy nebude znovu rozebrána.

**Účetní hodnota vlastního kapitálu/Závazky celkem ... (X<sub>3</sub>)**

Tato proměnná nebyla obsažena v původním Altmanově modelu z roku 1968, ale byla modifikována na tento tvar až v roce 1977 při tvorbě ZETA modelu. Původní tvar používal v čitateli tržní hodnotu vlastního kapitálu a byl vhodnější pro práci se společnostmi na akciovém trhu, přičemž ZETA model využívá účetní hodnotu vlastního kapitálu a je tak vhodnější pro soukromé firmy, na které se zaměřila tato práce.

Tato proměnná dává do poměru účetní hodnotu vlastního kapitálu a cizí zdroje, což je jinak obrácená hodnota ukazatele „dluh na vlastní kapitál“ (ang. Leverage). U této proměnné je stejně jako u proměnné č. 6 vcelku těžké rozhodnout o odhadovaném vlivu na bankrot společnosti. Je to dáno tím, že i když se zjednodušeně dá říci, že společnost s vysokým podílem vlastního kapitálu je stabilnější a tím i lepší, tak je obecně známo, že cizí kapitál je levnější a jeho nevyužití je tedy také nevýhodou.

Ideální hodnota se bude nacházet v kombinaci těchto zdrojů, ale vzhledem k faktu, že se tato hodnota nedá zobecnit, tak bude vycházeno ze zjednodušené myšlenky o větší výhodě při vlastním financování. Společnosti s vyšší hodnotou vlastního kapitálu na tom tedy budou lépe, a proto je odhadované znaménko této hodnoty záporné.

**Tržby/Aktiva celkem ... (X<sub>4</sub>)**

Následující proměnná je poslední proměnnou z původního Altmanova modelu vztahu tržby k celkovým aktivům. Tento ukazatel je také typickým pro klasickou finanční analýzu, kde se nazývá jako obrat aktiv a vyjadřuje, jak se zhodnocují aktiva v dané společnosti bez rozlišení zdrojů jejich krytí. Z této definice také vychází odhad znaménka proměnné X<sub>5</sub>. Vzhledem k faktu, že čím vyšší bude obrat aktiv, tím bude firma produktivnější, tak je vliv této proměnné na finanční situaci společnosti kladný a jeho odhadované znaménko je záporné.

**Log (Aktiva celkem/deflátor HDP) ... (X<sub>5</sub>)**

Tato proměnná je první proměnnou, která byla převzata z Ohlsonova modelu z roku 1980. Dává do poměru logaritmovaná aktiva firmy vůči HDP indexu cenové hladiny, což by mělo vypovídat o reálné velikosti firmy vzhledem k zemi, ve které operuje. Reálnou hodnotu zajišťuje použití HDP indexu cenové hladiny, který se

vypočítá jako poměr nominální a reálné hodnoty HDP. Tato hodnota se také nazývá jako deflátor HDP a měla by zajistit očištění hodnoty aktiv v tomto ukazateli o vliv změny cenové hladiny. Podle ČSÚ jsou hodnoty deflátoru pro potřebné roky 2012-2014: 1,5, 1,4, 2,5 % respektive 101,5, 101,4 a 102,5 % po přičtení 1.

Vzhledem k tomu, že tento ukazatel nepoměřuje čistě finanční hodnoty z dané společnosti, tak jej nelze zařadit do klasických ukazatelů finanční analýzy a jedná se spíše o specifický ukazatel, který kombinuje účetní data a makroekonomické údaje.

Z důvodů, které byly uvedeny už při výběru kritérií pro data set lze usoudit, že společnosti, které jsou menší (nižší hodnota aktiv), budou více náchylné k bankrotu a naopak. Proto je odhadované znaménko u této proměnné záporné.

### **Závazky celkem/Aktiva celkem ... ( $X_6$ )**

Šestá proměnná je klasickým ukazatelem finanční analýzy, který se nazývá ukazatel celkové zadluženosti. Tento ukazatel porovnává cizí zdroje vůči celkovému majetku firmy, čímž de facto vypovídá o tom, jaká část majetku firmy je financována věřiteli. Hodnoty této proměnné se nacházejí v intervalu  $X_6 \in <0,1>$ . U této proměnné je vcelku těžké přesně určit její vliv na bankrot společnosti, protože je obecně známo, že do určité míry je zadlužování společnosti přínosné a až od vysokých hodnot zadlužení má negativní efekt. Tento efekt spočívá v situaci, kdy věřitelé budou za své prostředky vyžadovat vyšší úrok, a to díky rizikové prémii, kterou si zaúčtují za riziko, že daná společnost vložené prostředky nedokáže splatit.

Přestože v menší míře má zadlužování pozitivní efekt, tak obecně bude tato proměnná brána jako proměnná s negativním vlivem a její odhadované znaménko je tedy kladné.

### **Krátkodobé cizí zdroje/Oběžná aktiva ... ( $X_7$ )**

Proměnná  $X_7$  porovnává krátkodobé cizí zdroje vůči oběžným aktivům, což je v zásadě převrácená hodnota ukazatele běžné likvidity, který se běžně využívá ve finanční analýze. Pro tuto převrácenou hodnotu není určen žádný doporučený interval, ale pokud je brán v úvahu fakt, že hodnota běžné likvidity by se měla pohybovat mezi  $<1,5; 2,5>$ , tak lze vyvodit, že pro proměnnou  $X_7$  tento doporučený interval bude mezi  $(0,4; 0,67)$ . Pokud tedy hodnota této proměnné bude spadat do určeného intervalu, tak lze očekávat kladný vliv (záporné znaménko) a pokud ne, tak vliv záporný (kladné znaménko).

### **Závazky celkem > Aktiva celkem ... ( $X_8$ )**

Pátý Ohlsonův ukazatel (8. pro tuto práci) nemá charakter ukazatele z klasické finanční analýzy, a dokonce nenabývá ani spojitých hodnot, jedná se tedy o ukazatel diskrétní. Sám Ohlson uvedl, že tento diskrétní ukazatel je určitou korekcí druhého ukazatele (6. v této práci), jelikož bere v úvahu specifickou situaci firmy, kdy cizí zdroje převýší celkový majetek firmy. V případě, že platí bilanční rovnice (aktiva=pasiva), může tato situace nastat pouze tehdy, pokud je vlastní kapitál záporný (např. velmi záporný výsledek hospodaření). Lze předpokládat, že tato proměnná bude v 90 % případů zanedbána, jelikož nabude nulovou hodnotu, ale pokud na-

stane případ, že nabude hodnotu 1, tak bude její vliv záporný a výsledné znaménko tedy kladné.

### **(Čistý zisk + odpisy) / Závazky celkem ... (X<sub>10</sub>)**

Další ukazatel je typickým ukazatelem finanční analýzy, který spadá do kategorie doplňkových ukazatelů, které ke svému výpočtu využívají cashflow. V tomto případě se jedná o poměr čistého zisku, navýšený o odpisy vůči celkovým závazkům společnosti, a vypovídá o tom, jak je daná společnost schopná splácet svoje dluhy. Při konstantní hodnotě čitatele by převrácená hodnota vypovídala o době splatnosti všech závazků dané společnosti. Je tedy zřetelné, že čím vyšší je tato hodnota, tím lepší vliv na společnost, a proto je odhadnuté znaménko záporné.

### **$(NI_t + NI_{t-1}) < 0$ ... (X<sub>11</sub>)**

Tato proměnná je druhým diskrétním ukazatelem, kterou Ohlson použil při tvorbě původního modelu z roku 1980. Vychází se z jednoduché myšlenky, že pokud v součtu za poslední dvě období je čistý zisk záporný, tak finanční situace společnosti není příznivá a bude náchylnější k bankrotu. Stejně jako u proměnné X<sub>8</sub> má tedy při hodnotě 1 negativní vliv a kladné výsledné znaménko.

### **$(NI_t + NI_{t-1}) / (|NI_t| + |NI_{t-1}|)$ ... (X<sub>12</sub>)**

Poslední ukazatel je opět specifický pro Ohlsonovu práci a nezapadá do žádné kategorie finanční analýzy. Při výpočtu tohoto ukazatele se pracuje s hodnotami čistého zisku dvou po sobě jdoucích období a určitým způsobem se jedná o zohlednění vývojové tendence tohoto zisku. V případě, že zisk roste ( $NI_t > NI_{t-1}$ ), tak dosahuje kladných hodnot a výsledné znaménko je záporné, ale pokud je situace opačná a zisk je menší než v minulém období, tak výsledný vliv hodnoty je záporný a znaménko proměnné je kladné. V třetí situaci, kdy jsou si čisté zisky z obou období rovny, tak výsledná hodnota nabývá hodnoty 0 a nemá žádný efekt.

Tab. 7 Odhadnuté vlivy a znaménka vysvětlujících proměnných

<b>Odhadnuté znaménko</b>	
X <sub>1</sub>	+
X <sub>2</sub>	-
X <sub>3</sub>	-
X <sub>4</sub>	-
X <sub>5</sub>	-
X <sub>6</sub>	+
X <sub>7</sub>	+
X <sub>8</sub>	+
X <sub>9</sub>	-
X <sub>10</sub>	-
X <sub>11</sub>	+
X <sub>12</sub>	±

### 3.3 Úprava dat

Je nutné přemýšlet nejen nad kritérii pro výběr data setu a vysvětlujícími proměnnými, ale také nad možnými úpravami dat samotných. Z literatury i praxe je známé, že obsáhlé databáze zahrnují i nepřiliš vhodné společnosti, co se týče kvality dat. Příkladem mohou být firmy, které nemají bankrotní status (vyhlášení bankrotu nebo likvidace), ale hodnoty jejich poměrových ukazatelů jsou mnohdy stejné nebo horší než u společností, které tento status bankrotu mají. Jinými slovy se jedná o situaci, kdy databáze nediskriminuje mezi firmami, od kterých sbírá data a neprobíhá zde žádná kontrola finančního stavu společností a jejich stavu. Z tohoto důvodu se do vzorku finančně zdravých společností dostávají i společnosti, které by již měly prohlásit bankrot nebo likvidaci, ale z nějakého důvodu tak ještě neučinily. Tímto důvodem může být např. finanční podpora od třetí strany (stát, mateřská společnost, ...), která má z určitého důvodu zájem o to, aby daná společnost zůstala v provozu. Pro představu této situace je možné přirovnat zmíněné společnosti k tzv. „zombie bankám“, což je termín používaný pro finanční instituce, které mají záporný vlastní kapitál, ale i nadále jsou uměle udržovány v provozu finančními dotacemi od státu. Tento termín se začal používat už v roce 1993 během Japonské krize.

Je tedy vhodné, aby byl takový data set o zmíněné společnosti očištěn. Pokud by tyto společnosti nadále zůstali v data setu i při modelování, tak by pravděpodobně negativně ovlivnily odhad ve prospěch zdravých společností. Respektive by model přiřazoval i finančně špatným společnostem status společnosti zdravé a velmi těžko by rozpoznával, která společnost by měla být označena jako bankrotní. Ve výsledném odhadu by tedy byla velmi nízká nebo dokonce nulová úspěšnost predikce bankrotu u zbankrotovaných firem.

Toto očištění spočívá v odstranění společností, které mají extrémně špatné hodnoty u jednoho nebo více vybraných poměrových ukazatelů, které s vysokou pravděpodobností vypovídají o špatném finančním zdraví dané firmy. Nejedná se tedy o vyhledání a přesun těchto společností do vzorku bankrotních společností, ale pouze o jejich vyřazení ze vzorku společností zdravých.

Na zhodnocení finančního zdraví podniku nestačí pouhá zadluženost daného subjektu, ale je potřeba se na něj dívat mimo jiné i z pohledu rentability, aktivity i likvidity. Z každé této oblasti byl vybrán alespoň jeden poměrový ukazatel, pomocí kterého bude zhodnoceno finanční zdraví (resp. finanční komplikace) společností ve vzorku nebankrotních firem.

Z pohledu rentability je vybrán ukazatel ROA ( $X_2$ ) neboli rentabilita aktiv. U tohoto ukazatele je doporučeno dosahovat nejméně 8 %, výnos bezrizikového aktiva se uvádí obecně kolem 3,5 % a po započítání inflace, rizikové marže a zisku je potřeba dosahovat alespoň této hodnoty, aby se vlastníkům vyplatilo investovat do dané společnosti, a nikoliv do jiných aktiv. Vzhledem k tomu, že není třeba vybírat pouze ideální společnosti, ale stačí pouze odstranit takové, které jsou nevhodné, tak je hraniční hodnota pro tento ukazatel posunuta na 5 %. Takováto hodnota by měla společnosti zajistit alespoň pokrytí bezrizikového výnosu a inflace.

Oblast aktivity je zastoupena proměnnou  $X_4$ , která udává rychlost obratu aktiv. Obecně se pro tento ukazatel doporučují hodnoty vyšší než 1, aby se tržby získané v daném roce vyrovnaly aktivům vloženým do podnikání. Avšak krátkodobý pokles této hodnoty pod 1 nemusí znamenat finanční tíseň společnosti, a proto je pro zmíněný data set minimální hodnota nastavena na 0,8.

Zadluženost společnosti popisuje proměnná  $X_6$ , přičemž doporučená hodnota v literatuře se uvádí 50 %. Kritická hodnota pro tuto práci je nastavena na hodnotu 0,8 vypovídající o situaci, kdy 80 % z celkového majetku společnosti je kryto dluhy. Taková společnost je očividně ve finančních potížích a zbytečně by znehodnocovala predikční schopnosti tvořeného modelu.

Z oblasti likvidity je vybrána převrácená hodnota běžné likvidity ( $X_7$ ). Takovýto tvar použil i Ohlson v původním modelu z roku 1980, a proto nebude pozměněn, ale je třeba určit alespoň rámcově hranici, ve které by se měly pohybovat zdravé společnosti. Pro ukazatel běžné likvidity se doporučují hodnoty (1,5;2,5) a jejich převrácená hodnota je tedy v intervalu  $\langle 0,4; 0,67 \rangle$ . Kritická hodnota pro získaný data set je opět mírnější než doporučených 0,67 a dosahuje hodnoty 0,8. Minimální hodnota je místo 0,4 nastavena na 0,2 a odpovídá hodnotě 5 u běžné likvidity. Kritická (maximální) hodnota vypovídá o příliš malé likviditě a minimální hodnota poukazuje na situaci, kdy má společnost zbytečně vysokou likviditu, která může negativně ovlivnit rentabilitu.

Poslední proměnnou pro tento výběr je proměnná  $X_{10}$ , která určuje, jakou část celkového dluhu společnosti by bylo možné splatit, pokud by věnovala veškerý svůj čistý zisk s odpisy na jeho splátku. Lépe představitelnou je převrácená hodnota této proměnné, která určuje dobu splatnosti celkového dluhu při konstantní výši zmíněné splátky. Pro účely této práce jsou vyřazeny takové společnosti, které by na splacení potřebovali minimálně 20 let, tudíž minimální hodnota této proměnné je 5 % (0,05).

Jednotlivé úpravy a na nich závislé změny v počtu společností v daném data setu budou uváděny v kapitole výsledky u jednotlivých modelů zvlášť.

## 4 Metodika

Tato kapitola se zabývá především rozebráním metod použitých k modelování a zhodnocení predikčních modelů. Závěrem této kapitoly je popsáno jakým způsobem tyto metody pro získání odpovědí na výzkumné otázky.

Z hlediska modelování predikčních bankrotních modelů existuje hned několik metod, přičemž ty hlavní byly zmíněny v literárním přehledu. Jednou z nejznámějších a dodnes velmi používanou metodou je logistická regrese. Existuje i několik nových způsobů, jako je např. BSM model nebo modelování pomocí neuronové sítě, ale již z dříve uvedených důvodů je pro tuto práci zvolen právě logit model.

### 4.1 Logistická regrese

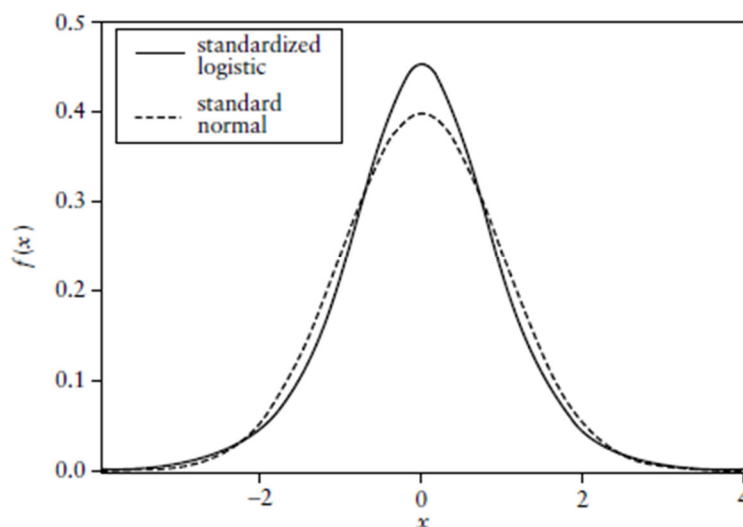
Pro účely této práce je tedy na místě, aby byl rozebrán postup sestavení logit modelu, aneb použití logistické regrese. Tento typ modelu je nejen široce citován v odborné literatuře, ale je vybrán i pro aplikaci na data set firem použitých v této práci, a proto je poskytnut jeho stručný rozbor.

Hlavní rozdíl logistické regrese oproti regresi lineární je forma závislých proměnných. Pro lineární regresi je vhodné použít pouze spojité závislé proměnné, které vyústí ve tvar  $Y = \beta_0 + \sum \beta x_i$ , ale pro nespojitě (diskrétní) proměnné je vhodné využít jiných forem modelu. Jednou oblastí diskrétních proměnných jsou proměnné binomické, které nabývají pouze hodnot 0 a 1. Pro jejich zpracování jsou vhodné modely binomické volby jako logit a probit. Pro využití v predikci bankrotu byl historicky vhodnější model logit, mimo jiné i proto, že je méně náročný na výpočet a není založen na natolik striktním předpokladu normálního rozdělení náhodné složky, ale principiálně mezi těmito dvěma modely není velký rozdíl. Vzhledem k tomu, že akademická obec dodnes využívá více formu logit modelu, která formovala i Ohlsonovo O-score, tak bude využita i pro tuto práci.

Jak ve svém díle napsal Valecký a Slivková (2012), tak lze předpokládat, že pravděpodobnost bankrotu  $i$ -té společnosti  $P_i = P(Y_i = 1)$  na základě vektoru  $x_i$ , lze vyjádřit pomocí funkce  $F(\beta; x_i)$ . Tato funkce je monotónně rostoucí  $F'(\beta; x_i) \geq 0$ , její definiční obor se pohybuje v intervalu  $(-\infty, +\infty)$  a obor hodnot v intervalu  $(0,1)$ . Pokud tedy platí, že  $F(-\infty) = 0$  a  $F(+\infty) = 1$ , tak obecnou funkci pravděpodobnosti úpadku společnosti lze zapsat jako  $P_i = F(\beta; x_i)$ . V této funkci  $\beta$  figuruje jako vektor parametrů  $(\beta_0, \beta_1, \beta_2, \dots, \beta_K)$  a  $x_i$  jako vektor vysvětlujících proměnných pro  $i$ -tou společnost  $(x_{1i}, x_{2i}, x_{3i}, \dots, x_{Ki})$ .

Heij et al. (2004) ve své práci zmiňují, že úspěšnost modelu nezávisí pouze na výběru vysvětlujících proměnných, ale také na tvaru distribuční funkce  $F$ . V praxi se poté často volí mezi standartním normálním rozdělením (model probit) a standartním logistickým rozdělením (model logit). Oba modely si jsou velmi podobné i v grafickém vyjádření jejich hustot, které jsou uvedeny v grafu níže:





Obr. 1 Porovnání rozdělení logit a probit (převzato z Heij et al., 2004)

Z tohoto grafu lze vyčíst, že logistické standardní rozdělení nabývá vyšších hodnot ve středních a koncových hodnotách (hodnoty  $x$  vzdálené od 0). Značnou výhodou logit modelu je možnost snadného analytického vyjádření kumulativní distribuční funkce  $F$  (v jejich díle autory označena jako  $\Lambda$ ):

$$\Lambda(t) = \int_{-\infty}^t \lambda(s) ds = \frac{e^t}{1 + e^t} = \frac{1}{1 + e^{-t}} \quad (11)$$

Na druhou stranu kumulativní distribuční funkce u modelu probit ( $F = \Phi$ ) by musela být číselně vypočtena pomocí aproximace integrálu. Rovnici lineární závislosti vysvětlované proměnné na proměnných vysvětlujících, kterou Ohlson pojmenoval  $Q$  (rovnice č. 2), zde autoři pojmenovali  $t$ . Vzhledem k faktu, že po dosazení dané rovnice  $t$  (nebo  $Q$ ) do finálního tvaru rovnice č. 11, je vypočtena pravděpodobnost bankrotu dané společnosti, tak bylo dosaženo stejného tvaru rovnice pro výpočet pravděpodobnosti, jakou uvedl Ohlson (rovnice č. 3), a to:

$$P = \left( \frac{1}{1 + e^{-Q}} \right)$$

Téma logistické regrese je velmi široké a cíleně byly vynechány některé části teoretického odvození, protože pro účely této práce by bylo toto odvození až zbytečně obsáhlé a je k dispozici v podrobném detailu v existující literatuře.

## 4.2 Metody hodnocení výsledných modelů

Po získání finálních modelů a jejich aplikací na data set je nutné, aby byla zhodnocena i kvalita daného modelu. Pro zhodnocení kvality modelu se nejčastěji používají dvě metody, a to křivka ROC a klasifikační tabulka. Obě zmíněné metody umožňují svým způsobem porovnat kvalitu i úspěšnost finálních modelů – křivka ROC graficky (AUC i početně) a klasifikační tabulka pomocí výsledné hodnoty určující úspěšnost klasifikace daného modelu (Valecký a Slivková, 2012).

### 4.2.1 Klasifikační tabulka

Klasifikační tabulka pojednává o úspěšnosti určitého modelu klasifikovat dané společnosti do skupiny finančně zdravých nebo ohrožených firem. V tabulce jsou tedy obsaženy údaje o četnosti správně a špatně klasifikovaných firem, které jsou rozděleny do čtyř hodnot, a to:

- *TP* („True positive“), neboli *dobře klasifikované* pozitivní případy
- *FP* („False positive“), neboli špatně klasifikované pozitivní případy
- *TN* („True negative“), neboli *dobře klasifikované* negativní případy
- *FN* („False negative“), neboli špatně klasifikované negativní případy

Tyto hodnoty jsou přímo udávány i ve výsledcích zmíněného programu Gretl, ve kterém budou modely odhadovány. Tyto tabulky mají obecně následující formu:

Tab. 8 Obecný tvar klasifikační tabulky

	Skutečné: 0	Skutečné: 1
Předpovězené: 0	<b>TN</b>	<b>FN</b>
Předpovězené: 1	<b>FP</b>	<b>TP</b>

Výsledná hodnota přesnosti predikce se vypočítá jako podíl správně klasifikovaných firem (součet hodnot na hlavní diagonále) k celkovému počtu firem. Matematicky se dá tento vztah znázornit jako  $\frac{TN+TP}{N}$ . Pochopitelně  $TN + TP$  je počet správně klasifikovaných firem a  $N$  zastupuje celkový počet firem v data setu.

K těmto údajům se často vážou i pojmy jako *senzitivita* (tj.  $\frac{TP}{TP+FN}$ ) a *specificita* (tj.  $\frac{TN}{TN+FP}$ ), ale vzhledem k tomu, že s těmito pojmy je pracováno v následující části, tak jsou vysvětleny níže.

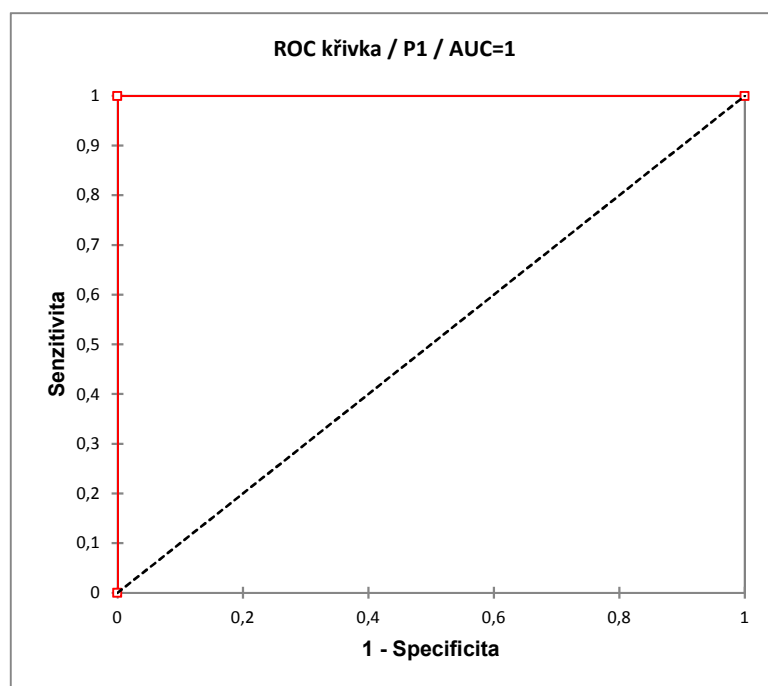
### 4.2.2 ROC křivka

Druhou možností, jak hodnotit diskriminační schopnost modelu je křivka ROC neboli „Receiver operating characteristic curve“, jak bývá označena v anglické literatuře. Tato křivka má na vertikální ose procentuálně vyjádřený počet firem, který byl správně klasifikován jako bankrotní (*TP*), vůči všem reálně zbankrotovaným společnostem ( $TP + FN$ ) a odborně se nazývá senzitivita. Jinými slovy tedy senzitivita udává schopnost modelu úspěšně klasifikovat bankrotní společnosti. Naopak

na horizontální ose je nanesena míra falešné positivity (1 - specificita), která vypovídá o poměru reálně zdravých společností, mylně označených modelem jako bankrotní, vzhledem ke všem zdravým firmám (tj.  $\frac{FP}{TN+FP}$ ) (Fawcett, 2006).

Tato křivka je tedy kombinací hodnot senzitivity a MFP (míra falešné positivity) při různých hladinách zlomu (v anglické literatuře nazývaných jako cut-off point). Tyto hladiny zlomu vyjadřují určitou hodnotu pravděpodobnosti, která rozhoduje, jestli je společnost klasifikována jako pozitivní nebo negativní. V případě predikce bankrotu je často využívána hladina zlomu v hodnotě 0,5, což znamená, že všechny společnosti, jejichž pravděpodobnost že zbankrotují je nižší než 0,5, jsou označeny za zdravé a pokud jsou vyšší než tato hodnota, tak jsou klasifikovány jako bankrotní.

Smyslem křivky ROC je tedy najít správnou hladinu zlomu, při kterém je zajištěný nejlepší poměr mezi maximalizací senzitivity a minimalizací hodnoty MFP. Je vhodné zmínit i možné extrémní situace na této křivce – křivka s levým horním rohem a diagonála. Pokud má křivka tvar zalomené přímky v levém horním rohu, tak je 100 % úspěšný a veškeré skutečné a odhadnuté hodnoty se shodují. Naopak pokud je křivkou diagonála, tak model má nulovou vypovídací schopnost a je naprosto nevhodný k predikci. Tyto extrémní situace jsou ilustrovány na následujících grafech.

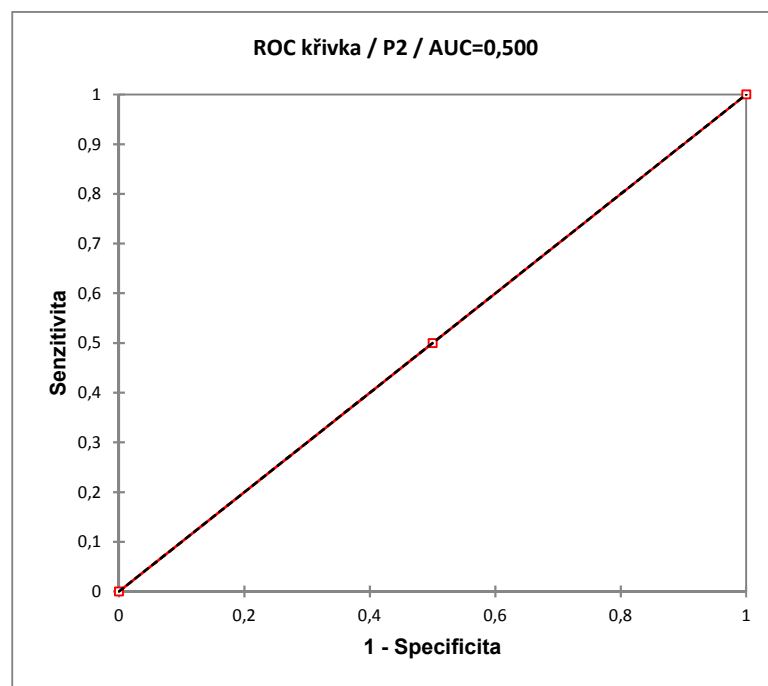


Obr. 2 ROC: dokonalý klasifikátor

Tento graf znázorňuje situaci, kdy model klasifikuje perfektně a naprosto se shodují predikce modelu se skutečností. Tato situace je možná pouze teoreticky nebo ve

zjednodušených případech, ale při modelování predikce reálných skutečností je téměř nemožné tohoto výsledku dosáhnout.

Na druhou stranu následující graf popisuje model, kde červená křivka znázorňuje druhou extrémní situaci, kdy model má stejnou predikční schopnost jako náhodné hádání.



Obr. 3 ROC: model nevhodný ke klasifikaci

K porovnání jednotlivých ROC křivek (jednotlivých modelů) se využívá ukazatel AUC (Area under curve), neboli plocha pod křivkou. Tento ukazatel je vyjádřen v jediné hodnotě a pohybuje se v intervalu  $<0; 1>$ , čímž je přímo vhodný k porovnávání. Modely vhodné k predikci by měly dosahovat hodnot vyšších než 0,5, přičemž vyšší hodnota tohoto ukazatele znamená kvalitnější model. Hodnota 0,5 udává situaci, kterou popisuje výše uvedený obrázek a jedná se o model, který je ekvivalentem náhodného hádání. Jinými slovy se jedná o 50% šanci, že bude vybraná společnost správně specifikována.

### 4.3 Spojitost s výzkumnými otázkami

Již při definování výzkumných otázek na začátku této práce, bylo zmíněno, že postup získání patřičných odpovědí bude blíže specifikován právě v části metodika. Vzhledem k tomu, že pro tento účel jsou využívány výše popsané metody, tak je možné rozebrat jejich spojitost.

U první výzkumné otázky byl uveden výraz predikční model, který je úspěšnější než náhodné hádání. V předchozí podkapitole byla tato skutečnost popsána extrémním případem křivky ROC (diagonála) a s ní spojeným koeficientem AUC ve

výši 0,5. Proto bude v následující kapitole kvalita modelu určována na základě grafického znázornění (ROC křivka) a hodnoty AUC. Kvalitní model musí mít hodnotu AUC vyšší než 0,5 a zároveň křivku, jenž je svojí většinou nad diagonálou grafu. Vzhledem k tomu, že grafické zhodnocení křivky ROC je poměrně subjektivní záležitostí, tak větší důraz je kladen na hodnotu AUC.

Druhá výzkumná otázka se již nezabývá pouze kvalitou modelu, ale přímo úspěšností predikce, která může být chápána ve třech kategoriích: celková úspěšnost predikce, úspěšnost predikce bankrotních společností a úspěšnost predikce zdravých společností. V závislosti na dříve zmíněné části o klasifikačních tabulkách jsou jednotlivé úspěšnosti vypočítány následovně:

- Celková úspěšnost:  $(TP + TN)/N$
- Úspěšnost zdravých:  $TN/(TN + FP)$
- Úspěšnost bankrotních:  $TP/(TP + FN)$

Z tohoto popisu je ale jasné, že jednotlivé modely se mohou lišit ve všech kategoriích a nebude možné lehce zjistit, který je vhodnější. Proto bude v těchto obtížně srovnatelných situacích přihlíženo opět ke koeficientu AUC a nově i k součtu hodnot senzitivity a specificity, na jejichž základě se dospěje ke vhodnému závěru.

Hodnoty pro jednotlivé kategorie jsou získány buď z původní hladiny zlomu (např. Ohlson, 1980 –  $P=0,5$ ) nebo z nové hladiny zlomu určené ROC analýzou.

## 5 Výsledky

V této kapitole jsou shrnuty veškeré praktické výsledky od sběru firemních údajů do potřebného data setu, přes jeho úpravu a zpracování, až po tvorbu samotných modelů a jejich porovnání s dalšími modifikacemi. V závěru této kapitoly jsou shrnuty výsledky práce a doporučení pro další výzkumy.

### 5.1 Výběr dat a tvorba data setu

Jak už bylo několikrát zmíněno, tak pro zrychlení a zpřehlednění údajů bylo pracováno s databází evropských společností a jejich výkazů Amadeus. Tato databáze umožňuje několikastupňové omezení hledaného vzorku až na výslednou hledanou množinu. Zmíněná množina pro tuto práci představuje vzorek společností po omezení (uplatnění omezujících kritérií), uvedených v podkapitole 3.1.

Použití zmíněných kritérií a mezivýsledky jsou uvedeny v tabulce č. 9, která přímo vyobrazuje omezující kritéria vyhledávání v programu Amadeus.

Tab. 9 Amadeus vyhledávání: Data set CZ\_2013 aktivní

	Mezivýsledky	Výsledky
1. Status: Aktivní	19 304 742	19 304 742
2. Oblast/Země/oblast v zemi: Česká Republika	488 882	476 436
3. NACE Rev. 2 hlavní sekce: F. Stavebnictví	2 488 426	37 312
4. Kategorie společností dle velikosti: Středně velké společnosti	2 789 348	5 310
Standardizovaná právní forma: Společnost s ručením omezeným, veřejná obchodní společnost, živnostník	20 148 639	4 745
6. Poslední rok účetnictví: 2015, 2014, 2013	19 430 948	<b>4 674</b>

V této tabulce jsou nejen vypsána všechna omezující kritéria, ale také mezivýsledky jednotlivých kroků vyhledávání. Sloupec mezivýsledků udává výsledky pro daný krok vyhledávání (dané omezení), přičemž následující sloupec s názvem výsledky udává vzorek společností, které vyhovují kumulovaným omezením. Například u čtvrtého kroku je použito omezení středně velkých společností, pro které vyhovuje 2 789 348 společností z celkové databáze programu Amadeus, kdežto pouze 5 310 společností vyhovuje kumulovanému omezení z kroků 1-4 (aktivní společnost + region ČR + sekce stavebnictví + středně velké společnosti). Kromě pátého kroku, který se zabývá formou podnikání (s.r.o., v.o.s., živnostník), je zahrnut i krok 6, který určuje poslední rok, ve kterém byly dostupné výkazy dané společnosti.

Porovnáním data setů pro roky 2015, 2014 a 2013 při omezení 1-5 bylo zjištěno, že rok 2013 je nejvhodnější pro tvorbu predikčního modelu, jelikož obsahoval

val nejvíce zbankrotovaných společností, a to 75. V roce 2015 to bylo 0 společností a v roce 2014 dvě společnosti. Data z roku 2013 jsou tedy nejdůležitější a je na nich formulován daný model.

Dále je nutné zavést šestý krok, který omezuje společnosti na takové, které měly dostupné výkazy minimálně ještě v roce 2013 nebo i v letech 2014 a 2015 (pro rok 2016 nejsou ještě dostupné údaje, a proto je rok 2015 brán jako současnost). Tento krok je zaveden hlavně kvůli menším firmám, které mohly vzniknout např. v roce 2012 a ještě není jisté, jestli budou provozuschopné, ale i přesto se už dostaly do databáze programu Amadeus. Takovéto společnosti většinou zanikají z jiných důvodů než zaběhlé provozuschopné firmy, tudíž by negativně ovlivnily predikční schopnost modelu a je nezbytné, aby ve finálním data setu nefigurovaly. Jejich vyloučení se provede výše zmíněným šestým kritériem, kdy se předpokládá, že společnost, která je v provozu alespoň jeden rok bez vyhlášení bankrotu nebo likvidace, je provozuschopná. S přihlédnutím k šestému omezení, je tedy 4 674 „zdravých“ společností, které pro rok 2013 vyhovují zmíněným kritériím.

Obdobně je získána i bankrotní část data setu, kde se ve vyhledávací matici pouze změní položka stavu „Aktivní“ na „Bankrot“, „Bankrot a rozpuštění spol.“, „Likvidace“ a „Likvidace a rozpuštění spol.“. Pro takové společnosti je účelné, aby právě rok 2013 byl posledním rokem, kdy má společnost dostupné výkazy, a proto je takto upraven i šestý krok. Výsledek vyhledávání je vyobrazen v následující tabulce.

Tab. 10 Amadeus vyhledávání: Data set CZ\_2013 bankrotní

	Mezivýsledky	Výsledky
1. Status: Likvidace, bankrot, Likvidace a rozpuštění spol., Bankrot a rozpuštění spol.	793 824	793 824
2. Oblast/Země/oblast v zemi: Česká Republika	488 882	7 846
3. NACE Rev. 2 hlavní sekce: F. Stavebnictví	2 488 426	1 034
4. Kategorie společností dle velikosti: Středně velké společnosti	2 789 348	110
5. Standardizovaná právní forma: Společnost s ručením omezeným, veřejná obchodní společnost, živnostník	20 148 639	92
6. Poslední rok účetnictví: 2013	4 263 132	75

Dalším krokem v programu Amadeus je nastavení požadovaných údajů z finančních výkazů jednotlivých společností. Tento program umožňuje získat téměř jakýkoliv údaj z klasických výkazů (rozvaha, výkaz zisku a ztrát, cash-flow), ale i několik nejdůležitějších poměrových ukazatelů. Bohužel žádný z potřebných ukazatelů v této databázi není, a proto jsou veškeré hledané údaje (s výjimkou deflátoru HDP) totožné s prvky poměrových ukazatelů obsažených v tabulce č. 6.

První úpravou této derivované databáze společností je promazání firem, kterým chybí část údajů, které brání výpočtu určitého ukazatele. Kupodivu je tento jev vcelku častý a je nutné databázi o takové společnosti očistit. Důvodem k tomuto promazání je fakt, že pokud budou zařazeny společnosti, kterým budou chybět některé poměrové ukazatele, tak budou chybně ovlivňovat celé určení vah poměrových ukazatelů. Po této úpravě zůstalo ve stažených databázích pouze 2264 zdravých společností (z původních 4764) a 28 bankrotních společností (z původních 75 firem). Tato úprava snížila počet firem v daných databázích o více než polovinu a je tedy zřetelné, že i když to je krok nezbytný, tak je vzhledem k počtu společností velmi drastický.

V takto upravených dočasných databázích je následně vypočítáno všech 12 poměrových ukazatelů pro veškeré společnosti v daném vzorku. Dalším krokem je vytvoření nové proměnné (sloupce), který představuje stav společnosti (zdravá vs. bankrotní). Tato proměnná je nazvána „Status“ a nabývá hodnoty 0 pro zdravé společnosti a 1 pro společnosti bankrotní. Tyto dvě očištěné databáze jsou spojeny a vzniká tak předčasné data set o 13 proměnných. Těchto třináct proměnných zastupuje jednu závislou proměnnou Status a 12 nezávislých proměnných, které jsou mimo jiné uvedeny v tabulce č. 6.

## 5.2 Modely pro české podniky (CZ)

V podkapitole o úpravě dat je zmíněno několik předpokladů a na nich založených úprav, které by měly vést k přesnějším modelům predikce, ale doposud nebyly tyto předpoklady potvrzeny na datech. Proto se následující podkapitola bude věnovat proměnám modelů, od původních neupravených data setů k novějším verzím, kde byly přijaty dříve zmíněné omezující předpoklady.

### 5.2.1 Model CZ\_1.1\_výchozí model

První model proto bude vytvořen na původním neupraveném data setu, u kterého se předpokládá špatná predikce bankrotních firem ve prospěch firem zdravých, a to z důvodu zahrnutí pochybných společností (typu „zombie bank“) do data setu.

Tento data set je složen z 2264 nebankrotních firem a 28 firem bankrotních. Jedinou úpravou dat v tomto data setu je odstranění firem s chybějícími údaji. Po načtení tohoto data setu o 2292 pozorováních, je tedy nutné vytvořit logit model pro binární vysvětlovanou proměnnou.

Při prvním odhadu modelu nejsou vylučovány proměnné, které jsou statisticky nevýznamné, a proto je běžné, že model je nutno o tyto proměnné očistit. V této práci se nejčastěji pracuje s daty, které obsahují pozorování v řádu tisíců, a tudíž pro testování byla zvolena hladina významnosti 1 %.

U prvního modelu je statisticky významný regresor pouze proměnná  $X_8$ , adjustovaný koeficient determinace ( $R^2$ ) dosahuje pouze hodnoty 0,106 a informační kritéria se pohybují okolo hodnoty 300 (Akaiikovo – 270,1; Schwarzovo – 344,6; Hannan-Quinnovo – 297,3). Nehledě na slabé ekonometrické charakteristiky tohoto modelu, je model slabý i co se týče predikce jako takové. Jeho celková úspěšnost



predikce sice dosahuje hodnoty 98,8 %, ale to je pouze díky nepoměru zdravých a bankrotujících společností. Při bližším pohledu je možno zjistit, že model sice správně odhadl veškeré zdravé společnosti, ale také mylně odhadl *veškeré* bankrotující společnosti. Bohužel, vzhledem k nevyváženosti data setu, se tento chybný odhad projevuje jako pouhé 1,2 % z celkového vzorku. Proto je nezbytné i u následujících modelů sledovat se zvýšenou pozorností úspěšnost predikce bankrotujících společností než úspěšnost predikce celkem. Pro lepší přehlednost charakteristik modelů, jsou výše zmíněné ekonometrické údaje i schopnost predikce uváděny v tabulkách pro každý model (i pro změny modelů). Jednotlivé vyobrazení kompletních modelů se všemi statistikami jsou uváděny v přílohách.

Tab. 11 Model CZ\_1.1

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,106	Celková	98,80 %
Akaikovo kritérium	270,1	Zdravé	100 %
Schwarzovo kritérium	344,6	Bankrotní	0 %
Hannan-Quinnovo kritérium	297,3		

Tímto prvotním odhadem ale práce na prvním modelu nekončí. Vzhledem ke špatným ekonometrickým charakteristikám modelu, bude otestováno, jestli nebude dosahovat lepších výsledků, po vypuštění statisticky nevýznamných proměnných. Toto odstranění proměnných z modelu nelze provést zároveň, protože při vypuštění jedné proměnné se změní p-hodnoty ostatních proměnných, a je tak nutné sledovat, které z nich je opravdu možno vypustit. Za použití sekvenční eliminace proměnných při podmínce, že p-hodnota je menší nebo rovna hodnotě 0,01, zbyly v modelu pouze dvě proměnné a konstanta. Těmito proměnnými jsou X<sub>8</sub> a X<sub>9</sub>. Po této úpravě model dosahuje následujících hodnot:

Tab. 12 Model CZ\_1.2\_výchozí model

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,115	Celková	98,70 %
Akaikovo kritérium	267,5	Zdravé	cca 100 %
Schwarzovo kritérium	284,7	Bankrotní	0 %
Hannan-Quinnovo kritérium	273,8		

Je viditelné, že po úpravě modelu se zlepšily ekonometrické charakteristiky, přičemž hodnota adjustovaného R<sup>2</sup> se zvýšila a hodnoty informačních kritérií poklesly. Na druhou stranu lze vyzorovat téměř zanedbatelný pokles v celkové úspěšnosti predikce, který je způsoben špatným odhadem jedné zdravé společnosti jako společnosti bankrotní (1 z 2268). Úprava modelu tedy vede k jeho zlepšení z hlediska ekonometrického, ale predikci nijak výrazně neovlivní.

Mohlo by se tedy zdát, že tvorba tohoto modelu je bezvýznamná, ale opak je pravdou. Nejen že tento model potvrdil předchozí domněnky a předpoklady o nut-

nosti další úpravy dat, v použitém data setu, ale stal se i pomyslným odrazovým můstkem pro modely, které budou z tohoto data setu nadále vytvářeny. Jeho důležitost spočívá hlavně v možnosti porovnání, přičemž na základě rozdílů s tímto modelem, bude rozhodováno o tom, jestli úpravy na data setu vedou k úspěšnějším predikcím.

### 5.2.2 Model CZ\_2.1\_dílčí modely

Při tvorbě předchozího modelu bylo potvrzeno, že model, vytvořený pouze na původním neupraveném data setu, nemá téměř žádnou vypovídací hodnotu. Vzhledem k tomu, že model nedokázal správně odhadnout ani jednu bankrotní společnost, tak se předpokládá, že je chyba v datech, kdy vzorek zdravých společností obsahuje společnosti se špatnou finanční situací, a tak model nedokáže rozeznat rozdíl mezi společnostmi zdravou a bankrotní. Díky značné nevyváženosti zdravých a bankrotních společností v data setu (ve prospěch zdravých), jsou proto všechny společnosti odhadnuty jako zdravé.

Řešením tohoto problému může být úprava „zdravé“ složky data setu, přičemž cílem je odstranit tyto nežádoucí „zdravé“ společnosti s velmi špatnými hodnotami vybraných proměnných. Touto úpravou by měla nastat situace, kdy model dokáže lépe rozeznat rozdíl mezi společnostmi zdravými a bankrotními, a tím zlepšit svůj odhad.

Při aplikaci jednotlivých omezení zmíněných v kapitole o úpravě data setu, je získáno několik verzí daného data setu. Tato omezení by samozřejmě mohla být provedena zaráz, ale tím by se nezjistilo, které omezení mělo nejlepší efekt na zlepšení predikce. Jinými slovy by se nezjistilo, která z oblastí dat u zdravých společností měla takové odchylky od normálního stavu, že to vychýlilo celkový odhad. Proto budou vytvořeny dílčí modely pro každý z upravených data setů. Výsledky dílčích modelů jsou uvedeny v tabulce č. 13.

Tab. 13 Model CZ\_2.1\_dílčí modely

<b>Charakteristiky</b>	<b>X<sub>2</sub></b>	<b>X<sub>4</sub></b>	<b>X<sub>6</sub></b>	<b>X<sub>7</sub></b>	<b>X<sub>10</sub></b>
Počet zdravých	928	1966	1647	1180	1478
R <sup>2</sup>	0,61	0,155	0,539	0,61	0,49
Akaikovo kritérium	98,4	248,6	131,2	103,74	142,1
Schwarzovo k.	161,6	321,4	196,3	170	211,2
H-Quinnovo k.	122,5	275,3	155,3	128,7	167,8
<b>Predikce</b>					
Celková	99.2 %	98,7 %	99,3 %	99,2 %	99,2 %
Zdravé	100 %	100 %	100 %	100 %	99,93 %
Bankrotní	71.4 %	7,14 %	60,7 %	64,29 %	60,71 %

V této tabulce jsou uvedeny dílčí neupravené modely, které byly vytvořeny na základě částečně omezených data setů. Je viditelné, že jednotlivá omezení mají odlišný dopad, jak na počet vyloučených společností ze vzorku „zdravých“ firem, tak na úspěšnost predikce.

Nejvyšší omezení je zaznamenáno při omezení proměnnou  $X_2$ , kde se vzorek zdravých společností zúžil na necelou tisícovku. Na druhou stranu se tímto zúžením velmi zpřesnil odhad bankrotních společností. K tomuto omezení se váže 71,4 % úspěšnost predikce bankrotních společností, což je samozřejmě vysoký rozdíl nejen vůči prvnímu modelu (úspěšnost predikce 0 %), ale také oproti úspěšnosti, která se váže ke zbylým omezením.

Naopak nejnižší omezení se váže k proměnné  $X_4$ , kdy zůstalo v data setu téměř dva tisíce společností. Na druhou stranu se dá očekávat, že takové laxní omezení bude mít pouze malý vliv na zpřesnění odhadu. Přesnost odhadu bankrotních společností pro toto omezení má hodnotu pouhých 7,14 %.

### 5.2.3 Model CZ\_2.2\_upravený model

Jelikož je žádaná co nejvyšší přesnost, tak není vhodné, aby bylo aplikováno pouze jedno omezení, ale postupně všechny tak, aby se omezily kritické hodnoty, které negativně ovlivňují predikci. Proto v následujícím modelu budou postupně po jednom aplikována omezení v pořadí podle jejich úspěšnosti predikce (z tabulky č. 13). Nejdříve tedy bude aplikováno omezení pro proměnnou  $X_2$ , ale vzhledem k tomu, že toto omezení a jeho výsledky jsou uvedeny ve zmíněné tabulce, tak přímo aplikujeme další, a to omezení provázané s proměnnou  $X_7$ .

Veškeré důležité výstupy z daného modelu jsou popsány v tabulce č. 14. Z této tabulky je jasné, že po přidání druhého omezení (k proměnné  $X_7$ ) se zvýšila přesnost odhadu u bankrotních firem o 17,9 %, tedy na hodnotu 89,3 %. Zlepšila se nejen predikce bankrotu ale i ekonometrické statistiky modelu – zvýšil se i adjustovaný koeficient determinace a všechna informační kritéria poklesla. Tento omezený vzorek společností se skládá z 576 společností

Tab. 14 Model CZ\_2.2\_úprava  $X_{2+7}$

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
$R^2$	0,765	Celková	99,3 %
Akaikovo kritérium	52,7	Zdravé	cca 100 %
Schwarzovo kritérium	109,3	Bankrotní	89,3 %
Hannan-Quinnovo kritérium	74,8		

Jako další je do modelu přidáno omezení vázající se k proměnné  $X_{10}$ . Po tomto omezení byl data set zúžen na 565 firem. Veškeré údaje, k dalšímu omezení a modelu, který se k němu váže, jsou opět zpracovány do následující tabulky.

Tab. 15 Model CZ\_2.2\_úprava  $X_{2+7+10}$ 

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
$R^2$	0,764	Celková	99,3 %
Akaikovo kritérium	52,7	Zdravé	cca 100 %
Schwarzovo kritérium	109	Bankrotní	89,3 %
Hannan-Quinnovo kritérium	74,7		

Z údajů uvedených v tabulce č. 15 je jasné, že na rozdíl od předchozích omezení, která zlepšovala predikci, tak omezení provázané k proměnné  $X_{10}$ , predikci nijak neovlivnilo. Neznamena to však, že by toto omezení nemělo smysl, ale pravděpodobně tato skutečnost poukazuje na fakt, že předchozí omezení, která se vázala k proměnným  $X_2$  a  $X_7$  už vyloučily většinu společností, které by byly předmětem daného omezení. Tuto myšlenku podpořil i fakt, že po omezení bylo z data setu vyloučeno pouze 11 zdravých společností. V této chvíli se nabízí otázka, jestli další dvě omezení budou mít dostatečný vliv na úspěšnost predikce, ale i kdyby jen pro ověření, tak je vhodné, aby byly predikční modely pro tato omezení vytvořeny a jejich data zhodnocena. Tento dílčí model se skládá celkem z 531 firem.

Tab. 16 Model\_CZ\_2.2\_úprava  $X_{2+7+10+6}$ 

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
$R^2$	0,784	Celková	99,4 %
Akaikovo kritérium	47,29	Zdravé	100 %
Schwarzovo kritérium	98,6	Bankrotní	89,3 %
Hannan-Quinnovo kritérium	67,3		

V tabulce č. 16 jsou vyobrazeny hodnoty pro dílčí model zahrnující i omezení provázané s proměnnou  $X_6$ , a i když zde není jasný žádný výrazný skok v úspěšnosti predikce, tak je znatelné zlepšení v ekonometrických statistikách. Vzhledem k tomu, že od druhého omezení se nezlepšila predikční schopnost pro bankrotní společnosti, tak je načase uvažovat o důvodu bankrotu u zbylých společností, které model nedokáže predikovat. Je možné, že jejich bankrot byl způsoben i jinými faktory, než jsou např. vysoká zadluženost a špatná likvidita. V případě, že by poskytnutý data set obsahoval robustní složku zbankrotovaných firem, tak by tyto společnosti mohly zastupovat pouhý zlomek z celku, ale v tomto data setu je bankrotních společností pouze 28 a 3 z nich není model schopen správně odhadnout. Avšak předtím, než budou přijata opatření pro zlepšení této situace, je nutné dokončit omezení i pro proměnnou  $X_4$ . Hodnoty z daného modelu jsou opět k vidění níže.

Tab. 17 Model\_CZ\_2.2\_všechny úpravy

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,7961	Celková	99,4 %
Akaikovo kritérium	44,03	Zdravé	100 %
Schwarzovo kritérium	94,66	Bankrotní	89,3 %
Hannan-Quinnovo kritérium	63,9		

Tabulka č. 17 obsahuje údaje z modelu, který zahrnuje všechna omezení, která byla dříve vybrána. Opět je viditelný pokrok v ekonometrických statistikách, ale i přesto je predikční schopnost modelu ve vztahu k bankrotním společnostem stále stejná. Jelikož je tato verze modelu finální, co se týče úpravy dat, tak je nutné model upravit do patřičné podoby z hlediska proměnných. Opět je tedy nutné ověřit statistickou významnost jednotlivých vysvětlujících proměnných a popřípadě je vyloučit z modelu. Údaje z upraveného modelu jsou uvedeny v tabulce č. 18.

Tab. 18 Model\_CZ\_2.2\_finální model

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,832	Celková	99,40 %
Akaikovo kritérium	36,22	Zdravé	100 %
Schwarzovo kritérium	57,31	Bankrotní	89,3 %
Hannan-Quinnovo kritérium	44,49	Hladina zlomu	<b>0,665</b>

<b>Proměnné</b>	<b>Koeficient</b>	<b>P-hodnoty</b>
Konstanta	-72,8	0,0093
X <sub>1</sub>	64,1	0,0113
X <sub>2</sub>	-109,2	0,0023
X <sub>3</sub>	3,7	0,0067
X <sub>7</sub>	81,5	0,0076

Již dříve bylo zmíněno, že při úpravě proměnných v modelu je žádoucí, aby byl model tvořen pro maximálně 1 % chybu prvního druhu (Type I error), ale to je za předpokladu, že změny vedoucí k této přesnosti jsou pro model vhodné a model jako takový zlepšují. Při úpravě výše zmíněného modelu byly ponechány proměnné X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub> a X<sub>7</sub>. Kromě proměnné X<sub>1</sub> všechny splňují podmínku pro 1 % chybu a proměnná X<sub>1</sub> má p-hodnotu pro t-test 0,0113, což značí hraniční případ. Při odebrání této proměnné bylo nutno postupně odebrat i další (změnili se jejich p-hodnoty) a v důsledku toho se postupně zhoršovaly nejen ekonometrické statistiky, ale i úspěšnost predikce. Proto bude v zájmu nejvyšší přesnosti modelu proměnná X<sub>1</sub> ponechána. Při této sestavě vysvětlujících proměnných je zastoupena část rentability, a to proměnnou X<sub>2</sub>, část poukazující na zadluženost společnosti

( $X_3$ ) a taky část zastupující běžnou likviditu (její převrácená hodnota) ve formě proměnné  $X_7$ .

Finální rovnice tohoto modelu má tedy tvar:

$$Q = -72 + 64,1X_1 - 109,2X_2 + 3,7X_3 + 81,5X_7 \quad (12)$$

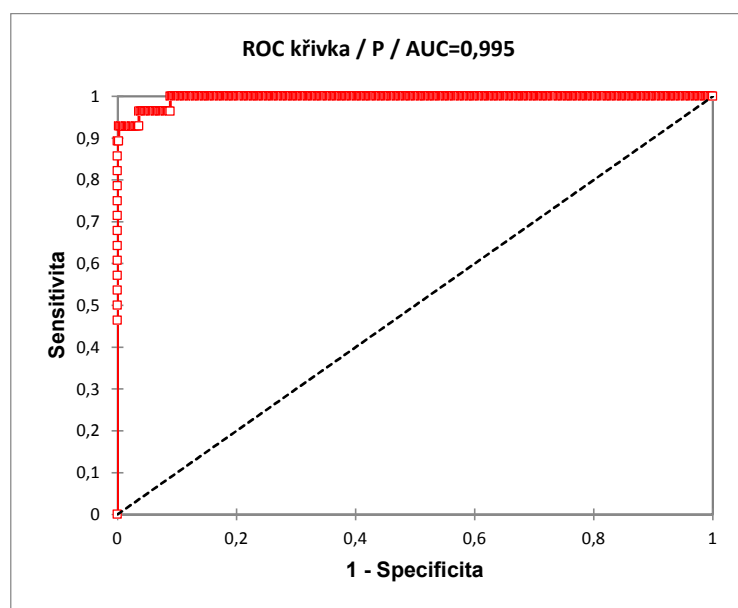
Tento model je ale nutné ještě verifikovat, a proto bude finální rovnice aplikována na nový data set (hold-out sample), kde bude zjištěna jeho úspěšnost na nových datech.

#### 5.2.4 ROC analýza původních dat

Tato podkapitola se zabývá verifikací posledního modelu CZ\_2.2\_finální model, jehož finální tvar je obsažen v rovnici č. 12. Ale než bude možné tuto verifikaci na nových datech (hold-out sample) provést, tak je nutné zajistit ROC analýzu tohoto modelu. V této analýze se jedná hlavně o nalezení správné hladiny zlomu (cut-off point), pro který je model nejúspěšnější ve své predikci. Program Gretl, ve kterém byly modely vytvářeny, tuto hladinu automaticky tvoří v oblasti, ve které model dosahuje nejvyšší přesnosti predikce, čímž se myslí poměr  $\frac{TP+TN}{N}$ . Např. zmíněný finální model z oblasti CZ, byl programem Gretl zalomen v  $P = 0,665$ , kterému odpovídají i uvedené úspěšnosti v tabulce č. 18. Jenže tento způsob určení hladiny zlomu nebere v potaz velikost vzorku pozitivních (bankrot) a negativních (zdravé) společností, a proto není příliš vhodný pro jeho aproximaci. Lepším způsobem, který je využíván i zahraničními autory je maximalizace hodnot specifity a senzitivity, potažmo maximalizace senzitivity a minimalizace hodnoty MFP (1 - specifita). Jak již bylo zmíněno v metodice, tak tyto hodnoty jsou vztaženy přímo ke svému vzorku pozitivních nebo negativních společností a nemělo by tak docházet k chybám při nerovnoměrném rozdělení data setu.

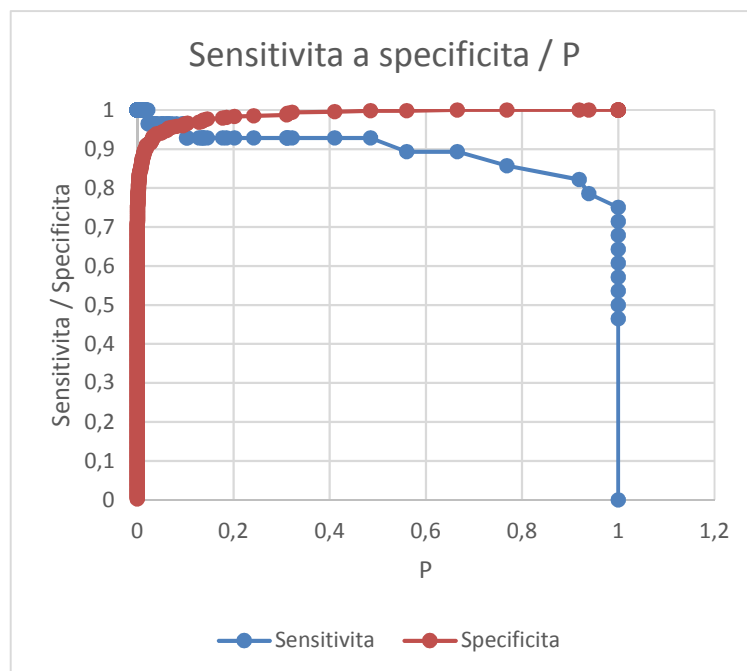
Vzhledem k tomu, že pro verifikaci těchto dat je hodnota zlomu nezbytná, tak je nutné nejdříve provést ROC analýzu modelu CZ\_2.5.5, která zjistí hladinu zlomu metodou specifity a senzitivity. Tato analýza je provedena v doplňku (add-on) programu Excel s názvem XLSTAT, kde spadá do kategorie „Survival analysis“. Součástí výsledné křivky ROC je i hodnota pod křivkou (AUC), která udává kvalitu modelu.

Konečná křivka ROC je uvedena na obrázku č. 4 a je značně podobná příkladu extrémů u této křivky (dokonalý klasifikátor) uvedené v metodice. Tento jev je způsoben hlavně tím, že křivka ROC byla použita na původní data, ze kterých byl vytvářen daný model. V pozdější části samotné verifikace se očekávají mnohem horší výsledky. Nehledě na to, lze na základě této křivky a hodnoty pod křivkou 0,955 (AUC) naznat, že tento model je opravdu kvalitní.



Obr. 4 ROC: CZ\_2.2\_finální model

Na základě grafického zhodnocení této křivky lze ale naznat, že hodnota hladiny zlomu pro tento model byla programem Gretl nastavena zbytečně vysoko (0,655). Z pohledu na tento graf je jasné, že hladina zlomu se bude pohybovat mnohem blíže nule, a to pravděpodobně mezi hodnotami 0 a 0,2.



Obr. 5 S+S: CZ\_2.2\_finální model

Pro ujasnění této hladiny je možné využít tohoto grafu, který udává přímo hodnoty senzitivity a specifickity ve dvou oddělených křivkách. V bodě, kde se střeťává křiv-

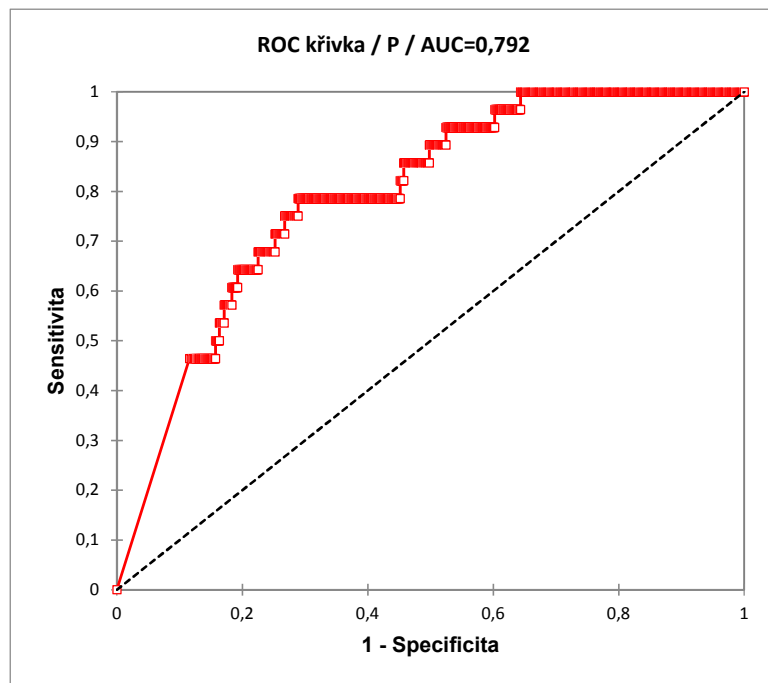
ka senzitivity a specificity, je nevhodnější hladina zlomu pro maximalizaci obou veličin, a tomuto bodu odpovídá hodnota  $P = 0,1$ . Je tedy značný rozdíl mezi hladinou zlomu, který je určen pomocí programu Gretl a tím, který je získán ROC analýzou. K tomuto bodu se také pojí odlišná klasifikační tabulka, než je v tabulce č. 18, a proto jsou nové údaje zaneseny do tabulky níže.

Tab. 19 Úspěšnost modelu CZ\_2.2\_finální při  $P = 0,1$

	Předpovězené		Úspěšnost	
	0	1		
Skutečné	0	457	17	
	1	1	27	
				Celková 96,41 %
				Zdravé 96,41 %
				Bankrotní 96,43 %

Z této tabulky lze lehce vyčíst, že i když se celková úspěšnost predikce lehce snížila, tak úspěšnost predikce pro bankrotní společnosti se značně zvýšila. Je to jasný důkaz o tom, že určení hladiny zlomu podle přesnosti (program Gretl) je nevhodné pro tento nerovnoměrně rozložený data set a je pro něj vhodnější způsob ROC analýzy.

I když je použití ROC analýzy na předchozí vzorek zajímavé, tak je faktem, že to je vzorek upravený a neodpovídá skutečnosti, kdy mnoho společností disponuje chybami popsány dříve. Proto je účelné, aby byl finální model aplikován na prvotní vzorek z prostředí ČR. Získáním těchto údajů lze zjistit, jak moc se liší hladiny zlomu v ideálních datech (upravené) a v datech běžných.



Obr. 6 ROC: CZ\_2.1\_výchozí model



Graf ROC pro zmíněný prvotní data set je vyobrazen na obr. č. 6. Z tohoto grafu je jasné, že finální model získaný na zúženém vzorku je sice méně, ale pořád kvalitní i pro data, kde se vyskytuje mnoho chyb. O tomto faktu vypovídá i hodnota AUC ve výši 0,792, která je nižší než u předchozích dat, ale pořád je značně vyšší než krajní hodnota 0,5. Už z tohoto grafického vyjádření je ale viditelné, že celý graf je vychýlený doprava a je zde značná falešná pozitivita. Toto zjištění, ale koresponduje s předpoklady, které vedly k upravení tohoto data setu na užší vzorek, ze kterého vznikl finální model CZ\_2.2. Předpokladem bylo, že v data setu se vyskytuje značné množství společností, které má natolik špatné finanční údaje, že jsou v ohrožení bankrotu, ale z neznámých důvodů jsou pořád označeny jako zdravé. Pro lepší představu jsou níže uvedeny údaje o úspěšnosti aplikovaného modelu při hladinách 0,1 a 0,665.

Tab. 20 Úspěšnost modelu CZ\_2.1 při  $P = 0,1$ 

		Předpovězené		Úspěšnost predikce	
		0	1		
Skutečné	0	898	1359	Celkem	0,404814
	1	1	27	Zdravé	0,397873
				Bankrotní	0,964286

Tab. 21 Úspěšnost modelu CZ\_2.1 při  $P = 0,665$ 

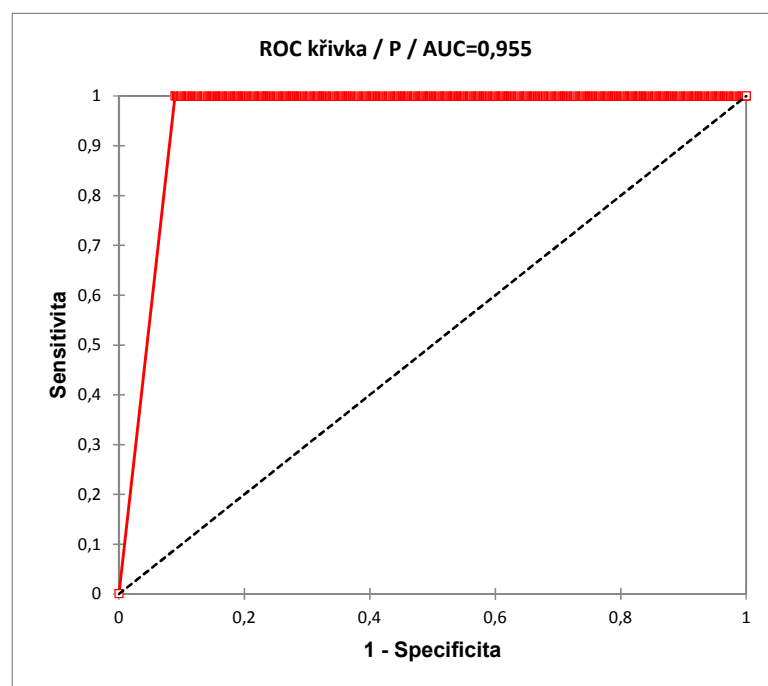
		Předpovězené		Úspěšnost predikce	
		0	1		
Skutečné	0	1099	1158	Celkem	0,49190372
	1	3	25	Zdravé	0,48692955
				Bankrotní	0,89285714

V tabulce č. 20 a 21 jsou k vidění úspěšnosti modelů při zlomových hladinách 0,1 a 0,665. Na první pohled se zdá, že je výhodnější využít hladinu 0,665, ale vzhledem k tomu, že ideální hladina zlomu pro tento vzorek se blíží 1, tak toto rozhodnutí není tak jednoduché. Pokud je brán v úvahu předchozí předpoklad o falešně pozitivních společnostech, tak značnou část špatné predikce tvoří špatné společnosti s označením jako zdravé. Není tedy jednoduché přesně určit, jaký počet zdravých společností je opravdu v pořádku a jaký ne. Jelikož ve finálním vzorku pro tvorbu modelu po očištění zůstalo cca 500 zdravých společností, tak je pravděpodobně lepší predikce 898 zdravých společností při hladině 0,1 než 1099 při hladině 0,665.

### 5.2.5 Verifikace modelu CZ

Ani tyto údaje nejsou jasným potvrzením úspěšnosti daného modelu, protože jsou pořád získávány ze stejného data setu, a proto je potřeba tento model otestovat na úplně nových datech. Bohužel, časově nejbližší data z let 2012 nebo 2014, jsou na-

prosto nedostatečná. V roce 2012 je pouze jedna bankrotní společnost s dostupnými daty a v roce 2014 jsou tyto společnosti dvě. Jakékoliv výstupy z těchto dat mohou být brány pouze s odstupem a mají malou vypovídací schopnost. I tak je vhodné, aby takový výstup, i kdyby jen pro referenci, byl vytvořen. Proto byl zvolen rok 2014 s 2 bankrotními a 1658 zdravými společnostmi jako testovací data set. Po aplikaci ROC analýzy vznikl následující graf:



Obr. 7 ROC: CZ\_Verifikace\_upravený model

Tento graf sice poskytuje informace o tom, že model je téměř dokonalým klasifikátorem s plochou pod křivkou ve výši 0,955, ale vzhledem k malému počtu bankrotních společností, nemají výsledky téměř žádnou vypovídací schopnost, a proto nebudou vytvářeny žádné úsudky. Jedinou informací, kterou je možno brát vážně je chování zdravých společností po aplikaci modelu. V následujících tabulkách jsou uvedeny hodnoty úspěšnosti modelu při hladinách 0,1 a 0,665.

Tab. 22 První verifikace modelu CZ při  $P = 0,1$

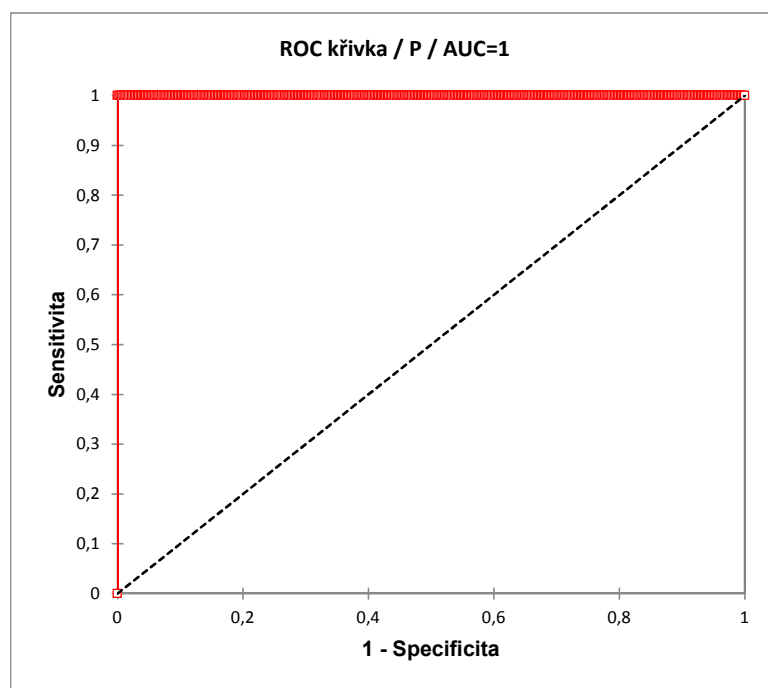
		Předpovězené		Úspěšnost predikce	
		0	1		
Skutečné	0	786	872	Celkem	0,474699
	1	0	2	Zdravé	0,474065
				Bankrotní	1

Tab. 23 Úspěšnost modelu CZ na nových datech při  $P=0,665$ 

		Předpovězené		Úspěšnost predikce	
		0	1		
Skutečné	0	944	714	Celkem	0,569880
	1	0	2	Zdravé	0,569361
				Bankrotní	1

Z těchto tabulek nelze jednoznačně vyvodit, že by byl model úspěšný v predikci zdravých společností (v oblasti TN), ale za předpokladu existence špatných dat, by se měla situace po vyčištění data setu výrazně zlepšit. Proto byl tento nový data set očištěn stejnými podmínkami, jako data sety předchozí (omezení u  $X_{2,4,6,7,10}$ ).

Tento nový očištěný data set byl podroben ROC analýze a stejně jako na neupravených datech, dosahuje 100 % predikce bankrotních společností, ale výrazný pokrok byl proveden právě v predikci společností zdravých. Při zlomu v hladině 0,1 byly zdravé společnosti predikovány s 97,3 % přesností a při hladině 0,665 s téměř 100 %. Rozdíl mezi těmito hladinami by měl být viditelný na data setu s více bankrotními společnostmi, kdy by při vyšších hodnotách zlomu, měl stoupat počet falešně negativních společností. Proto nelze tvrdit, že u tohoto modelu je vhodnější použít hladinu zlomu na základě celkové přesnosti (0,665) nebo na základě ROC analýzy (0,1) a bylo by pro toto tvrzení nutno získat kvalitnější data set. Pro představu je uveden i graf druhé verifikace.



Obr. 8 ROC\_CZ\_Verifikace\_finální model

Tento graf popisuje použitý model, jako perfektní klasifikátor vzhledem k použitému data setu, a proto je hodnota AUC ve výši 1. Tato situace je samozřejmě v praxi téměř nemožná, což přidává další důvod k tomu, aby byla tato verifika-

ce modelu na novém data setu brána pouze jako reference, ne jako důkaz nebo ověření úspěšnosti modelu.

### 5.2.6 Souhrn modelu CZ

V této podkapitole se práce zaměřila na získaný data set a jeho zpracování do úspěšného modelu. Ale hned po prvním predikčním modelu, který byl na původním data setu vytvořen, nastal problém. Zjistilo se, že údaje společností z databáze Amadeus nejsou ideální pro tvorbu predikčního modelu a obsahují mnoho společností, které neodpovídají charakteru běžné „zdravé“ společnosti.

Proto byla potřeba určitá selekce zdravé složky data setu, které bylo provedeno na základě zvolení kritických hodnot u vybraných proměnných. Veškeré společnosti, které tyto hodnoty překročily, byly z data setu vyloučeny a model se dál zdokonaloval. Postupným přidáváním omezení bylo zjištěno, že nejvyšší vliv na zlepšení predikce má proměnná  $X_2$  a dá se tudíž předpokládat, že tato položka byla nejčastěji porušována společnostmi ve zdravém vzorku.

Při upravování data setu se zjistilo, že daný model má další problém – jeho bankrotní část je příliš malá na to, aby mohl vzniknout opravdu kvalitní a obecněji použitelný model. Tato skutečnost se projevila tím, že pouhé tři společnosti, které pravděpodobně zbankrotovali díky jiným důvodům, než které se projevují ve výkazech společnosti, zastupovaly více než 10 % celkové neúspěšnost odhadu v daném modelu. Je jasné, že zahrnutí těchto společností nelze vyloučit, ale při robustnějším vzorku bankrotních společností by pravděpodobně zaujaly značně nižší zlomek z celkové úspěšnosti predikce.

Toto dilema vyřešila aplikace ROC analýzy, při které se zjistilo, že je výhodnější využívat hladinu zlomu, která byla určena na základě senzitivity a specificity daného vzorku, nikoliv na základě celkové přesnosti predikce. Úspěšnost predikce daného modelu i předpokladu, že je vhodnější do predikce zahrnout ROC analýzu, byla testována na verifikačním data setu. I když tento nový data set umožnil potvrdit myšlenku, že v databázi Amadeus se pravidelně vyskytují zdravé společnosti s bankrotními hodnotami ve finančních výkazech, tak neposkytl odpověď na verifikaci úspěšnosti. Za touto situací stojí jediný fakt, a to nedostatek bankrotních společností ve vybraném vzorku.

Tento problém není lehce řešitelný a při pokusu o jeho překonání je nutné přijmout drastické změny. Existují dvě možnosti, které tento problém řeší, ale obě určitým způsobem porušují předpoklady, ze kterých se vycházelo pro tvorbu současného predikčního modelu. Tímto řešením může být rozšíření vzorku z pohledu místa nebo času. Jinými slovy lze rozšířit vzorek tak, že v něm bude zahrnuto více států nebo budou získána data z minulých let a vytvoří se pouze jeden vzorek s daty z mnoha let, ve kterém se ale ztratí právě vliv času. Vzhledem k tomu, že při tvorbě tohoto modelu se vybrala data z jednoho roku a z ČR právě proto, aby byl model přesnější a vystihoval lépe podmínky dané lokality a období, tak obě tyto změny jsou porušením daných předpokladů.

Ovšem, při promyšlení těchto možností, je mnohem výhodnější možnost rozšíření lokality než času, a to ze dvou důvodů. Toto rozšíření lokality může být pro-

vedeno například „pouze“ pro celou EU, o které se dá tvrdit, že má podobné hospodářské podmínky. Druhým důvodem je fakt, že modifikace starších modelů se tvoří právě kvůli faktu, že čím je testovaná společnost časově dále od období, na kterém byl vytvořen model, tím klesá přesnost predikce. Pokud by tedy bylo přijato rozšíření z hlediska času, tak by byl vytvořen určitý průměr z těchto období, který vůbec nemusí odpovídat současnosti. Z těchto důvodů je vybrána možnost rozšíření data setu z hlediska lokality a dále se tato práce zaměří na získání a zpracování data setu společností z celé EU a jeho zpracování do úspěšného predikčního modelu.

### 5.3 Modely pro data podniků z EU

Aby bylo možné tvořit a upravovat predikční modely pro oblast Evropské unie, tak je nutné opět získat data set tomu odpovídající. Je tedy zopakován celý proces výběru dat v databázi Amadeus, pouze s tím rozdílem, že se zruší omezení na zemi původu. Databáze Amadeus totiž získává data pouze evropských zemí, a proto pouze stačí, když se zruší omezení ve vztahu k lokalitě. Vzhledem k tomu, že je třeba vybírat data ze zemí, které mají přibližně stejné prostředí, tak budou primárně vybrány země EU nebo jejich partneři.

Ale ani v této oblasti není vyhledávání údajů zcela přehledné a jednoduché. Obecně je známo, že ne všechny evropské státy jsou členem Evropské unie, i když s ní mají značný kontakt a dobré vztahy. Jsou to např. státy jako Norsko, Island nebo Švýcarsko. I přesto, že tyto státy nejsou členem EU, tak mají natolik silné obchodní a politické vztahy, že jejich národní prostředí je velmi podobné tomu v EU, a budou proto také zahrnuty do evropského data setu.

Dále existuje několik států, které sice ještě v Evropské unii nefigurují jako plnohodnotní členové, ale je s nimi počítáno jako s budoucími členy. Těmito státy jsou Černá Hora a Srbsko, které spadají do kategorie kandidátských zemí, ale i Bosna a Hercegovina, která je zatím uváděna pouze jako potenciální kandidátská země. Tyto státy budou také přidány do data setu, protože i pro přijetí do seznamu kandidátských zemí je nutné splnit určitá nejen hospodářská kritéria EU a dá se tedy tvrdit, že mají podobné prostředí jako zbytek EU.

Na druhou stranu je ale značný problém v samotné databázi Amadeus, protože se sice snaží získávat data z téměř všech evropských zemí, ale některá zde z neznámých důvodů nefigurují. Takovými zeměmi jsou např. Dánsko nebo Spojené království.

Pro lepší přehlednost jsou všechny evropské státy, které byly vybrány pro tvorbu následného data setu uvedeny v tabulce č. 24, která je uvedena na následující straně.

Výsledkem hledání pro tato území a výše zadaná kritéria je soubor 235 774 aktivních a 3749 bankrotních společností. Detailní výstup tohoto hledání i s mezivýsledky u jednotlivých kritérií je k nalezení v přílohách. Při zpracování tohoto obrovského množství společností se opět setkáváme s nedostatkem dané databáze. Stejně jako u dat z českého prostředí, se i zde nachází vysoký počet společ-

ností, které nemají veškeré údaje potřebné pro výpočet poměrových ukazatelů obsažených ve vysvětlujících proměnných, a je tedy nutné tuto databázi očistit.

Tab. 24 Použité země EU

<b>Evropské země použité v data setu</b>	
Belgie	Maďarsko
Bosna a Hercegovina	Malta
Bulharsko	Německo
Černá hora	Nizozemsko
ČR	Norsko
Dánsko	Polsko
Estonsko	Portugalsko
Finsko	Rakousko
Francie	Rumunsko
Chorvatsko	Řecko
Irsko	Slovensko
Island	Slovinsko
Itálie	Spojené království
Kypr	Španělsko
Litva	Srbsko
Lotyšsko	Švédsko
Lucembursko	Švýcarsko

### 5.3.1 Model EU\_1.1\_výchozí

Po očištění společností bez všech údajů, bylo dosaženo vzorku 90 630 aktivních a 1264 bankrotních firem. Ale jak už bylo zmíněno výše, tak rozsah chybějících údajů v dané databázi je opravdu značný, a to dokonce natolik, že z data setu vypadly celé země. Například všechny společnosti ze Spojeného království měly neznámé tržby, které jsou potřeba pro výpočet proměnné  $X_4$ . Dalším příkladem mohou být litevské společnosti, kdy ani jedna z nich neměla uvedené údaje o amortizaci a depreciaci (nutné k výpočtu proměnné  $X_{10}$ ). Těchto příkladů by mohlo být více, ale faktem zůstává, že touto očištěnou bylo vyřazeno z data setu 11 evropských zemí z důvodu chybějících údajů. Pro lepší představu je na další straně opět uvedena tabulka, ve které jsou veškeré země figurující v data setu po očištění chybějících údajů.

Tento upravený data set se dá považovat za první použitelný data set z této oblasti, ale i před vytvořením modelu, který se k němu váže, je možné předpovědět, že tento model nebude velice úspěšný, co se predikce týče. Už na modelu z českého prostředí bylo zjištěno, že v databázi se vyskytují společnosti, které sice

mají status aktivní, ale jejich finanční výkazy mnohdy vypovídají o horší situaci než u společností bankrotních.

Tab. 25 Použité země EU (upraveno)

<b>Evropské země použité v data setu</b>	
Belgie	Nizozemsko
Bosna a Hercegovina	Norsko
Bulharsko	Polsko
ČR	Portugalsko
Estonsko	Rakousko
Finsko	Rumunsko
Francie	Slovensko
Chorvatsko	Slovinsko
Itálie	Španělsko
Lucembursko	Srbsko
Maďarsko	Švédsko
Německo	

Příkladem důvodu k takovému stavu je již dříve zmíněný zombie bank atd., proto nebude znovu rozebírán. Proto i když je téměř jisté, že tento model bude špatný, co se predikce týče, tak je nutné jej vytvořit minimálně pro vytvoření základu, se kterým budou porovnávány pozdější modely. Jednotlivé údaje z tohoto modelu jsou sepsány v následující tabulce.

Tab. 26 Model EU\_1.1

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,129	Celková	98,6 %
Akaikovo kritérium	11576,15	Zdravé	cca 100 %
Schwarzovo kritérium	11698,6	Bankrotní	cca 0 %
Hannan-Quinnovo kritérium	11613,46		

Tato tabulka potvrzuje veškerá předchozí očekávání o špatné predikční schopnosti tohoto modelu. Nejen že má model velmi špatné ekonometrické charakteristiky, ale i jeho úspěšnost predikce u bankrotních společností je téměř nulová. Na druhou stranu byl tento model vytvořen ze všech dvanácti vysvětlujících proměnných a nebyl upravován v závislosti na jejich významnosti.

Tab. 27 Model EU\_1.2\_výchozí

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,129	Celková	98,6 %
Akaikovo kritérium	11577,2	Zdravé	cca 100 %
Schwarzovo kritérium	11633,7	Bankrotní	cca 0 %
Hannan-Quinnovo kritérium	11594,4		

I když se po úpravě lehce zlepšily popisná kritéria, tak model zůstává stejný, co se týče adjustovaného koeficientu determinace i úspěšnosti predikce. Detailní výstupy k původní i upravené verzi jsou opět uvedeny v přílohách. V této fázi je nutné přemýšlet o zlepšení tohoto modelu a stejně jako u modelu z českého prostředí se nabízí úprava vstupních dat. Tato úprava je uplatněna ze stejného důvodu jako dříve – v kategorii středně velkých podniků se často nacházejí takové, jež mají velmi špatné finanční výkazy, ale i přesto jsou označeny jako aktivní. Proto stejnou úpravu při stejných podmínkách filtrování jsou aplikovány i na data set pro EU.

### 5.3.2 Model EU\_2.1\_finální

Jelikož při modelování predikce pro oblast ČR již bylo vyzkoušeno, že všechna navržená omezení mají určitý kladný vliv na predikční schopnost daného modelu, tak nebudou testovány zvlášť a přidávány jednotlivě, ale hned budou aplikována všechna omezení a pro tento data set bude vytvářen následující model.

Po úpravě původního data setu a vyfiltrování nevhodných společností, byl získán soubor společností, složený z 20 603 zdravých a 1260 bankrotních firem. Nej důležitější statistiky tohoto modelu jsou uvedeny v následující tabulce a plný výstup je uveden opět v přílohách.

Tab. 28 Model EU\_2.1

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,76	Celková	98,9 %
Akaikovo kritérium	2313,6	Zdravé	cca 100 %
Schwarzovo kritérium	2409,5	Bankrotní	82,06 %
Hannan-Quinnovo kritérium	2344,8		

Z tabulky je jasně viditelný dramatický růst, jak kvality modelu z hlediska ekonometrického, tak z hlediska úspěšnosti predikce. Adjustovaný koeficient determinace vzrostl o cca 0,63 a zároveň se snížila hodnota informačních kritérií zhruba o 9 tisíc bodů, což v obou případech znamená dramatický růst kvality modelu. Z hlediska predikce se model také velice zlepšil, a to z nulové hodnoty až na solidních 82,06 % při predikci bankrotních společností. Tento výstup je z modelu, který není upravený z hlediska vysvětlujících proměnných, a proto je opět nutné zjistit statistickou významnost  $X_1$ - $X_{12}$ .

Tab. 29 Model EU\_2.2\_finální

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,76	Celková	98,9 %
Akaikovo kritérium	2307,95	Zdravé	cca 100 %
Schwarzovo kritérium	2379,89	Bankrotní	82,06 %
Hannan-Quinnovo kritérium	2331,39	Hladina zlomu	<b>0,5</b>



Tabulka č. 29 popisuje již upravený model z hlediska významnosti proměnných. Model se sice jen lehce zlepšil v ekonometrických charakteristikách a schopnost predikce zůstala stejná, ale na druhou stranu byly zjištěny statisticky významné proměnné, které by tvořily finální rovnici tohoto modelu.

Na rozdíl od modelů z prostředí ČR, obsahují evropské modely více statisticky významných proměnných. V tomto modelu jsou významné proměnné  $X_1$ ,  $X_3$ ,  $X_5$ ,  $X_6$ ,  $X_7$ ,  $X_9$ ,  $X_{11}$  a  $X_{12}$ . Vyřazeny z modelu byly jen proměnné  $X_2$ ,  $X_4$ ,  $X_8$  a  $X_{10}$ . Model tedy bere v úvahu proměnné ze všech důležitých oblastí, jako je rentabilita ( $X_9$ ), zadluženost společnosti ( $X_3$  a  $X_6$ ), likvidita ( $X_7$ ), ale i specifické znaky společností, jako je jejich velikost ( $X_5$ ), vývoj čistého zisku ( $X_{11}$  a  $X_{12}$ ) a poměrná velikost čistého pracovního kapitálu vzhledem k celkovým aktivům ( $X_1$ ).

Bylo tedy dosaženo velmi dobrého výsledku u úspěšnosti predikce, kdy model téměř nedělá chyby při rozpoznání zdravých společností a u bankrotních je dokáže odhadnout ve více než 80 % případů. Otázkou zůstává, jestli není možné i tento model nějakým způsobem zlepšit.

Nabízí se myšlenka úpravy bankrotní složky data setu podobným způsobem jako u zdravé složky. Přesněji řečeno je pravděpodobné, že stejně jako u modelu pro ČR, tak i zde budou figurovat společnosti, které jsou označeny jako bankrotní, i přesto že jejich finanční výkazy tomu neodpovídají. U takových společností se pravděpodobně nejedná o klasický bankrot vzhledem k nedostatku finančních prostředků, vysokého zadlužení atd., ale spíše se bude jednat o speciální případy vzhledem k psychickému rozpoložení majitelů (psychický kolaps, rozhodnutí pohybovat se v jiné činnosti atd.). Je tedy potřebné najít způsob, jak tyto společnosti identifikovat, následně je vyloučit z data setu a sledovat, jestli se predikční schopnost modelulepší nebo nikoliv. Rovnice této modifikace evropského modelu je následující:

$$Q = -24,85 + 9,94 * X_1 + 1,78 * X_3 + 0,68 * X_5 + 16,26 * X_6 + 11,93 * X_7 - 12,35 * X_9 - 1,39 * X_{11} - 2,57 * X_{12} \quad (13)$$

### 5.3.3 Model EU\_3.1\_upravený

K této identifikaci lze použít stejná kritéria pro očištění zdravého vzorku od nevhodných společností. Provedení by spočívalo ve vyfiltrování bankrotních společností, které nesplňují tato kritéria pro zdravé společnosti a zbydou společnosti bankrotní, které tyto kritéria splňují. Jinými slovy jsou získány bankrotní společnosti, které ani v jedné z vybraných kategorií nepřekračují kritické hodnoty.

Po tomto očištění je zjištěno, že existuje 60 společností, které splňují zároveň všechna kritéria, a tudíž jsou nevhodné pro zařazení do data setu bankrotních společností. Pro zlepšení daného modelu, ale nestačí pouhé zlepšení predikce, ale je nutné sledovat, jestli se predikcelepší více, než jen při vyloučení libovolných 60 společností. Pokud by měl být uveden příklad na předchozím modelu, kde je ve vzorku bankrotních společností 1260 celkem společností (1034 odhadnutých

správně, 226 špatně), tak pouhým odstraněním 60 špatně odhadnutých společností se predikční schopnost zvýší na 86,17 %. Takováto úprava by byla pouhým zúžením vzorku a nedá se o ní tvrdit, že zlepšil model jako takový. Pro opravdu vhodnou úpravu je nutné, aby se predikční schopnost zvýšila ještě více, jelikož se předpokládá, že tyto odstraněné společnosti vychylovaly odhad ve prospěch zdravých společností a po jejich odstranění, odstraníme i toto vychýlení.

Evropský model pro upravenou bankrotní složku data setu je uveden v tabulce č. 30.

Tab. 30 Model EU\_3.1

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,856	Celková	99,4 %
Akaikovo kritérium	1341,37	Zdravé	cca 100 %
Schwarzovo kritérium	1437,25	Bankrotní	89,33 %
Hannan-Quinnovo kritérium	1372,61		

Tato tabulka obsahuje údaje nového modelu a jelikož ekonometrické statistiky i úspěšnost predikce jsou lepší než v modelu EU\_2.2, tak je jasné, že předchozí úprava byla správná. Úspěšnost predikce bankrotních společností dokonce překonala zmíněných 86,17 % a dosahuje hodnoty 89,33 %, což vypovídá o tom, že odstraněné společnosti opravdu způsobovaly vychýlení v predikci.

Je možné spekulovat o tom, jestli by další zúžení data setu (bankrotní části) nezlepšilo úspěšnost predikce ještě více, ale další zásahy jsou dalším omezením bankrotní části, která je i tak mnohem menší než část zdravá. Proto je zúžení bankrotní části provedeno maximálně ve výši kritických hodnot použitých i na zdravou část data setu. Uživatel finálního modelu bude mít sám možnost posoudit podle vypočítané hodnoty, jestli je společnost na pokraji intervalu, který znamená bankrot a podle toho jí také přisoudit požadovanou pozornost.

I přes fakt, že byly v poslední verzi modelu získány skvělé výsledky, tak je nutné, aby byl model opět očištěn o statisticky nevýznamné proměnné a dospělo se tak k finální rovnici daného modelu.

Tab. 31 Model EU\_3.2\_upravený

<b>Ekonometrické charakteristiky</b>	<b>Hodnoty</b>	<b>Typ predikce</b>	<b>Úspěšnost</b>
R <sup>2</sup>	0,856	Celková	99,4 %
Akaikovo kritérium	1337,8	Zdravé	cca 100 %
Schwarzovo kritérium	1417,7	Bankrotní	89,42 %
Hannan-Quinnovo kritérium	1363,83	Hladina zlomu	<b>0,498</b>

Model je nakonec očištěn jen o proměnné X4 a X9. Proměnná X1 také nesplňuje striktní požadavek na maximálně 1 % chybu (p-hodnota nižší než 0,01), protože její p-hodnota dosahuje hodnotu 0,011, ale vzhledem k tomu, že po jejím odstraně-

ní došlo ke zhoršení úspěšnosti predikce i ekonometrických charakteristik, tak bude v modelu ponechána.

Finální rovnice pro evropský model má následující formu:

$$Q = -23,46 + 3,88X_1 - 4,43X_2 + 3,59X_3 + 0,59X_5 + 21,72X_6 + 7,98X_7 - 16,72X_{10} - 1,95X_{11} - 2,95X_{12} \quad (14)$$

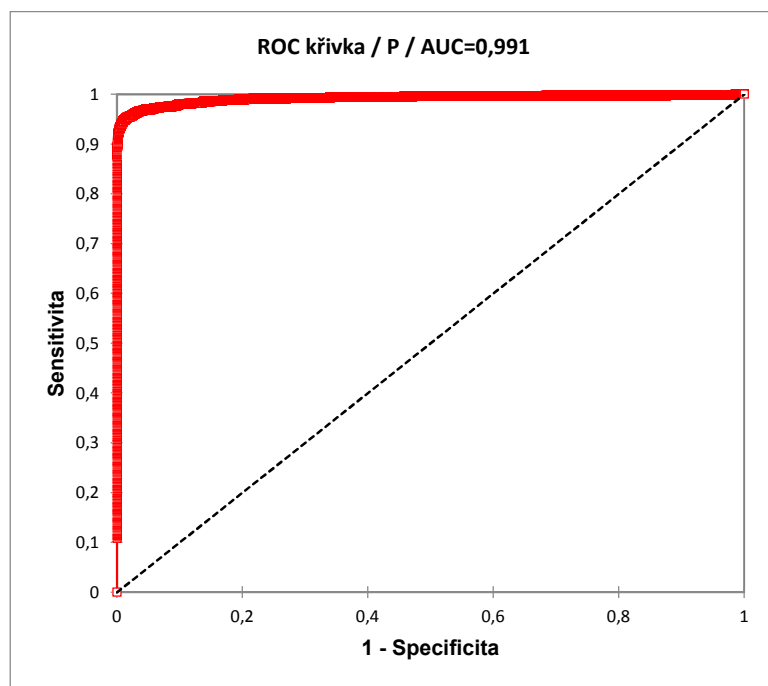
Tento model je nutné aplikovat na jiná data, než na kterých byl vytvořen, a proto je nutné získat kontrolní data set (hold-out sample) a vytvořený model otestovat v praxi. Ale ještě před samotnou verifikací je vhodné na původní data aplikovat ROC analýzu a zjistit nejvhodnější hladiny zlomu.

### 5.3.4 ROC analýza na původních datech

Stejně jako u aplikace ROC analýzy na modelu CZ je i tady nutné zjistit hladinu zlomu na základě specifity a senzitivity, která bere v potaz i velikost jednotlivých složek data setu. Tuto hladinu je nutno porovnat s původní hladinou zlomu, která byla určena na základě celkové přesnosti modelu programem Gretl.

Avšak oproti ROC analýze v modelu CZ je zde nutné postupovat odlišně. Díky třem verzím modelu (neupravená, upravená nebankrotní data, upravena i bankrotní data) vyvstala otázka, který z modelů bude nejvhodnější. Při modelování vycházel samozřejmě nejlépe poslední model se všemi upravenými daty, ale to neznamená, že bude dosahovat nejlepších výsledků i při ROC analýze nebo pozdější verifikaci. Tato pochybnost vznikla na základě obavy o to, jestli poslední úpravou nebyl model příliš vychýlen ve prospěch bankrotních společností. Dalším důvodem je značný rozdíl mezi tvarem druhé a třetí modifikace tohoto modelu, které se liší nejen v koeficientech, ale i v proměnných. První neupravený model nemá téměř žádnou vypovídací schopnost, a proto bude porovnávána pouze druhá a třetí modifikace tohoto modelu. U obou modifikací budou zohledněny a otestovány i rozdíly mezi hladinami zlomu určené podle přesnosti (0,5 pro druhou modifikaci a 0,498 pro třetí) a následně získané hladiny zlomu z ROC analýzy. I když je možné, že na těchto původních datech nebudou mezi modifikacemi modelů takové rozdíly, tak v pozdější verifikaci na nových datech mohou být rozdíly značné.

Je vhodnější začít třetí modifikací modelu, která je vytvořena na nejvíce upravených datech. Analýza pro třetí modifikaci bude odlišena znaky „-3“, např. ROC: Model EU\_3.2-3. Analýza pro druhou modifikaci potom logicky získá označení „-2“ (ROC: Model EU\_3.2-2).



Obr. 9 ROC: Model EU 3.2-3

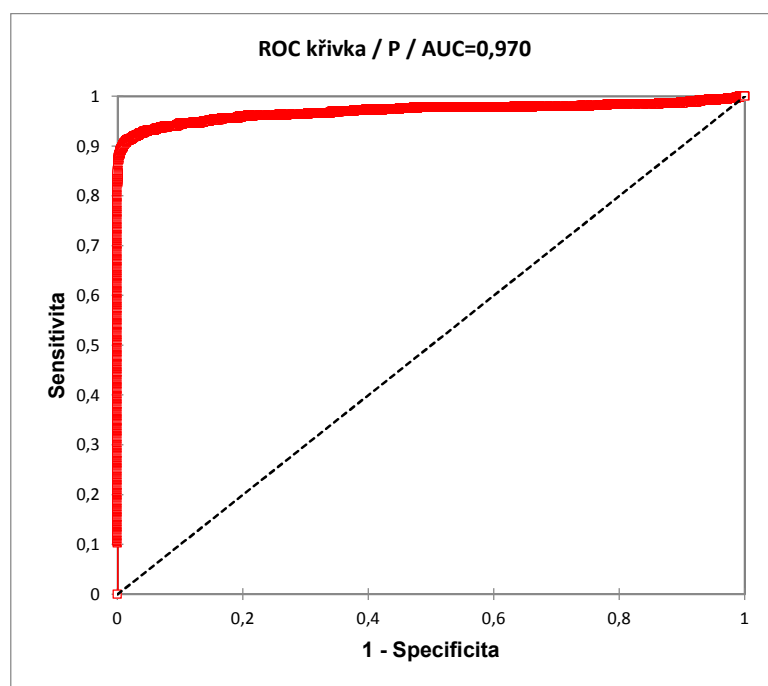
Tato křivka, patřící k modelu 3.2, vypovídá o téměř dokonalé klasifikaci dat pomocí třetí modifikace modelu EU. K této téměř dokonalé křivce se pojí i velmi vysoká hodnota  $AUC=0,991$ . Na základě ROC analýzy byla opět určena rozdílná hladina zlomu na základě maximalizace senzitivity a specificity, a to v hodnotě  $P=0,108$ . Porovnání úspěšnosti predikce mezi původní a novou hladinou zlomu je vypsána v následující tabulce. Vzhledem k vysokému množství ROC analýz tohoto modelu jsou následující porovnání vyobrazena pouze v procentuální úspěšnosti bez klasifikační tabulky:

Tab. 32 Model EU\_3.2-3, pro  $P=0,108$  a  $0,498$ 

<b>P = 0,108</b>		<b>P = 0,498</b>	
Celková	98,40 %	Celková	99,40 %
Zdravé	98,60 %	Zdravé	cca 100 %
Bankrotní	95,16 %	Bankrotní	89,42 %

Tyto výsledky, stejně jako u modelu CZ, poukazují na větší vhodnost hladiny zlomu získanou ROC analýzou. Sice díky němu o pár procent klesla predikční schopnost zdravých společností, ale na druhou stranu, predikční schopnosti pro společnosti bankrotní, vzrostla na 95,16 %.

Dále tento model bude aplikován na data s upravenými nebankrotními společnostmi, ale bez upravených společností bankrotních.



Obr. 10 ROC: Model EU\_2.2-3

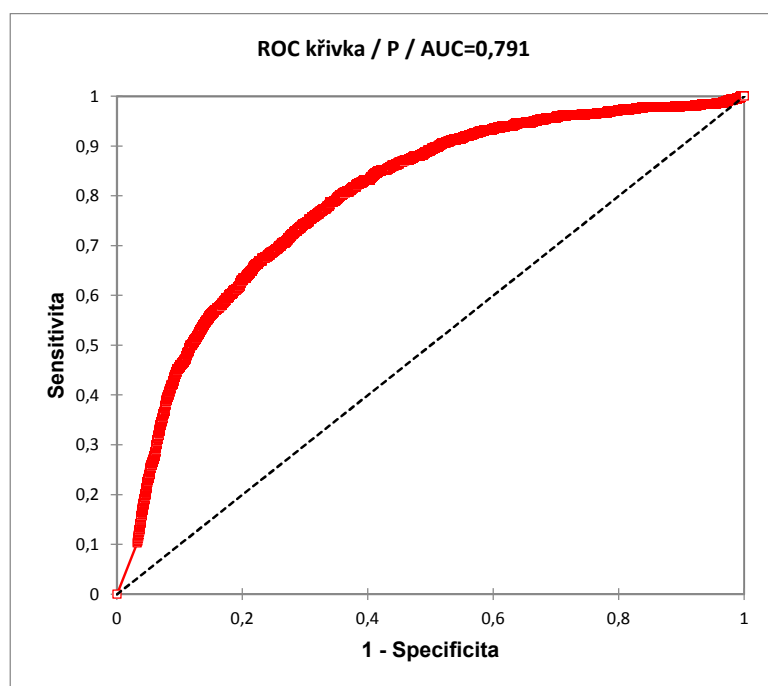
Tento graf poukazuje na skutečnost, že model vytvořený na upravených datech, nedokáže v tomto data setu vysvětlit větší část společností než na data setu předchozím, což se pojí i s horším sklonem křivky ROC a nižší hodnotou AUC (0,97). Předpokládá se, že tyto špatně klasifikované společnosti jsou právě ony upravené (vymazané) bankrotní společnosti, které měly finanční výkazy odpovídající společnostem zdravým. V takovém případě by měla značně klesnout predikční schopnost pro bankrotní společnosti, což lze jednoduše ověřit v následující tabulce:

Tab. 33 Model EU\_2.2-3, pro  $P = 0,108$  a  $0,498$ 

<b>P = 0,108</b>		<b>P = 0,498</b>	
Celková	98,20 %	Celková	99,10 %
Zdravé	98,60 %	Zdravé	cca 100 %
Bankrotní	91,03 %	Bankrotní	85,15 %

Tato tabulka potvrzuje výše uvedený předpoklad, přičemž lze jasně vidět sníženou schopnost predikce bankrotních společností, za jinak nezměněné úspěšnosti predikce společností zdravých. Opět tyto údaje potvrzují, že je vhodnější zvolit hladinu zlomu na základě ROC analýzy (0,108).

Dále je vhodné tuto analýzu a porovnání provést i na neupravených datech původního data setu, kde se čeká nejnižší predikční schopnost. Tento fakt je dán tím, že z tohoto data setu bylo mnoho společností odstraněno hlavně kvůli jejich zavádějícím finančním výkazům, na kterých je celý model vybudován.



Obr. 11 ROC: Model EU\_1.2-3

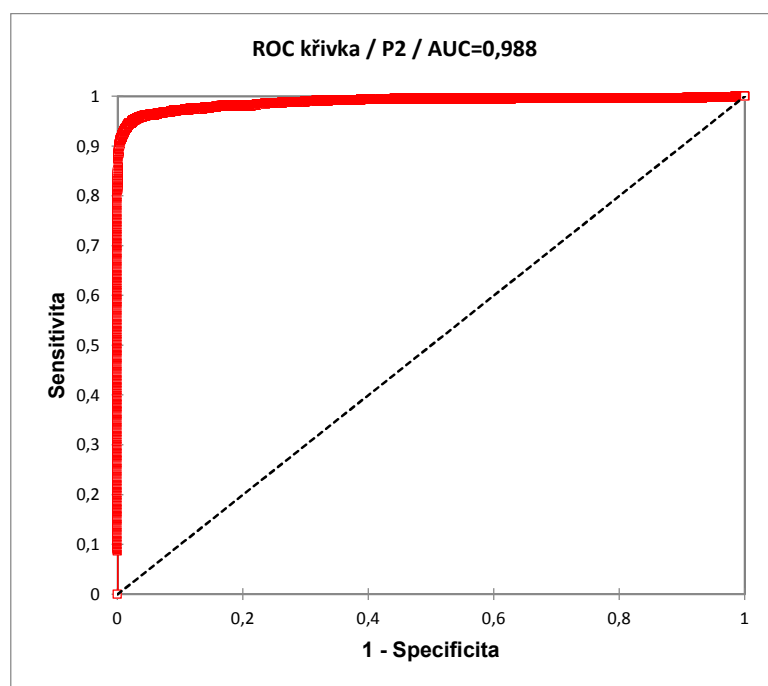
I když lze z tohoto grafu vyčíst mnoho nepřesností, tak pořád dosahuje velmi dobrých hodnot, kdy nejen hodnotou AUC (0,791), ale i grafickým znázorněním poukazuje na dobrou predikční schopnost daného modelu. Srovnání úspěšnosti predikce, pro obě možné hladiny zlomu, jsou v následující tabulce:

Tab. 34 Model EU\_1.2-3, pro  $P=0,108$  a  $0,498$ 

<b>P = 0,108</b>		<b>P = 0,498</b>	
Celková	47,60 %	Celková	58,00 %
Zdravé	47,04 %	Zdravé	57,60 %
Bankrotní	91,03 %	Bankrotní	85,15 %

V této tabulce je poznat dříve zmíněný propad v úspěšnosti predikce zdravých společností, ale je třeba se na tuto situaci dívat s nadhledem, protože tato nepřesná klasifikace „zdravých“ firem nastává právě u společností, které mají natolik špatné finanční výkazy, že byly z data setu vyloučeny pro značnou podobnost bankrotním společností. Proto, i když je celková úspěšnost pro  $P=0,498$  lepší, nelze jednoznačně tvrdit, že je pro predikci vhodnější, protože v jejím souboru „zdravých“ společností figurují právě i výše zmíněné společnosti se špatnými výsledky.

Stejný proces je teď nutné zopakovat také pro model EU\_2.2 (a odpovídající rovnici Q), aby bylo otestováno, jestli model EU\_3.2 nebyl moc vychýlen odstraněním části bankrotních společností.



Obr. 12 ROC: Model EU\_3.2-2

Graf č. 12 poukazuje stejně jako u třetí modifikace modelu na téměř dokonalou klasifikaci dat pomocí druhé modifikace modelu EU. Dokonce i úspěšnosti predikce tohoto modelu jsou velmi podobné, jak je vidět z následující tabulky:

Tab. 35 Model EU\_3.2-2, pro P=0,092 a 0,5

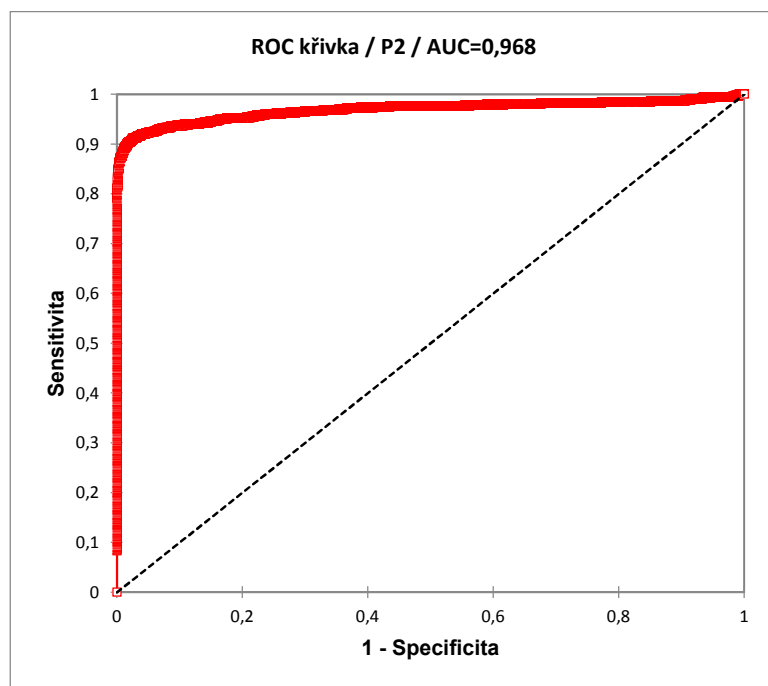
<b>P = 0,092</b>		<b>P = 0,5</b>	
Celková	97,30 %	Celková	99,20 %
Zdravé	97,44 %	Zdravé	cca 100 %
Bankrotní	95,25 %	Bankrotní	86,16 %

Další část se věnuje druhé modifikaci modelu EU na data s neupravenými bankrotními společnostmi, což je de facto přesně data set, na kterém tento model vznikl.

Opět byla zaznamenána situace, kdy nastalo snížení úspěšnosti klasifikace, ale i tak zůstává model velmi dobrý a ukazatel AUC je větší než 0,9 (0,968). Opět se očekává, že toto snížení bylo díky bankrotní části data setu, o čemž se lze přesvědčit v následující tabulce:

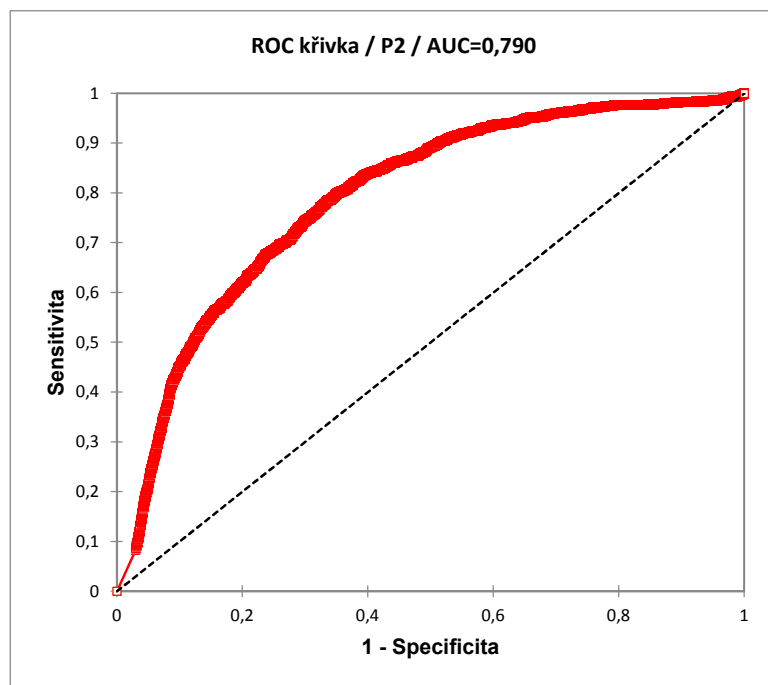
Tab. 36 Model EU\_2.2-2, pro P=0,092 a 0,5

<b>P = 0,092</b>		<b>P = 0,5</b>	
Celková	97,10 %	Celková	98,90 %
Zdravé	97,44 %	Zdravé	cca 100 %
Bankrotní	91,10 %	Bankrotní	82,06 %



Obr. 13 ROC: Model EU\_2.2-2

Jako poslední je opět zanalyzována původní neupravená část data setu, která je zobrazena v následujícím grafu:



Obr. 14 ROC: Model EU\_1.2-2



Stejně jako u třetí modifikace tohoto modelu, tak i druhá modifikace dosahuje i na původních datech velmi dobrých výsledků. Tento fakt lze ověřit i v následující tabulce, která uvádí úspěšnost predikce tohoto modelu na daném data setu.

Tab. 37 Model EU\_1.2-2, při P=0,092 a 0,5

<b>P = 0,092</b>		<b>P = 0,5</b>	
Celková	47,70 %	Celková	62,40 %
Zdravé	47,10 %	Zdravé	62,14 %
Bankrotní	91,03 %	Bankrotní	82,06 %

Tato tabulka poukazuje na situaci, kdy opět nelze lehce říci, jestli je vhodnější použít hladinu zlomu z programu Gretl nebo z ROC analýzy. Vzhledem k faktu, že v tomto data setu se vyskytuje mnoho „zdravých“ společností se špatnými výkazy, tak bude nadále předpokládáno, že hladina zlomu zvolená pomocí analýzy ROC je vhodnější.

Na druhou stranu rozhodnutí, zdali je vhodnější využít model EU\_3.2 nebo EU\_2.2 nelze z této části rozhodnout, jelikož jejich výsledky jsou na původních datech téměř totožné. Proto bude o vhodnosti rozhodnuto až po jejich aplikaci na nová data v části verifikace modelu.

### 5.3.5 Verifikace modelu EU

Již začátkem podkapitoly bylo zmíněno, že bude vytvořena nová verze evropského data setu, na kterém bude otestována úspěšnost finální rovnice daného modelu. Tento nový data set je opět získán z databáze Amadeus, a to téměř při stejných podmínkách, jako u data setu původního. Jediným rozdílem je rok, ve kterém byly údaje získány, přičemž při tvorbě modelů se vycházelo z dat roku 2013 a pro verifikaci jsou použity data z roku 2014.

Nový získaný data set pro rok 2014 se skládá z 185 338 aktivních společností a z 2674 společností bankrotních. Tyto údaje jsou samozřejmě před jakýmkoliv očištěním od nežádoucích jevů a jednotlivé kroky a mezivýsledky vyhledávání jsou opět uvedeny v přílohách. V tuto chvíli je potřeba věnovat zvýšenou pozornost tomu, jaké země se v novém data setu nacházejí. Před úpravou v něm byly obsaženy veškeré země, jako v data setu původním (tabulka č. 19), s výjimkou Černé Hory, ale po odstranění všech společností s chybějícími hodnotami, se v něm vyskytuje už pouhých 24 států. Tento počet je dokonce o jedno vyšší, než v původním data setu (tabulka č. 20), ale je nutné zjistit, jaké země jsou shodné a jaké ne. Veškeré země použité v tomto očištěném data setu jsou v tabulce č. 38 na následující straně.

Z této tabulky lze vyčíst, že prvním rozdílem, oproti původnímu data setu, je absence dvou zemí, a to Portugalska a Španělska. Druhým rozdílem je přidání tří nových zemí, oproti původnímu data setu, kterými je Island, Řecko a Lotyšsko. Vzhledem k tomu, že tato změna není velmi výrazná, tak je předpokládáno, že tento nový data set je velmi podobnou aktualizací původního evropského data setu.

Tab. 38 Evropské země k verifikaci modelu EU (očištěno)

<b>Evropské země použité v data setu</b>	
Belgie	Maďarsko
Bosna a Hercegovina	Německo
Bulharsko	Nizozemsko
ČR	Norsko
Estonsko	Polsko
Finsko	Rakousko
Francie	Rumunsko
Chorvatsko	Řecko
Island	Slovensko
Itálie	Slovinsko
Lotyšsko	Srbsko
Lucembursko	Švédsko

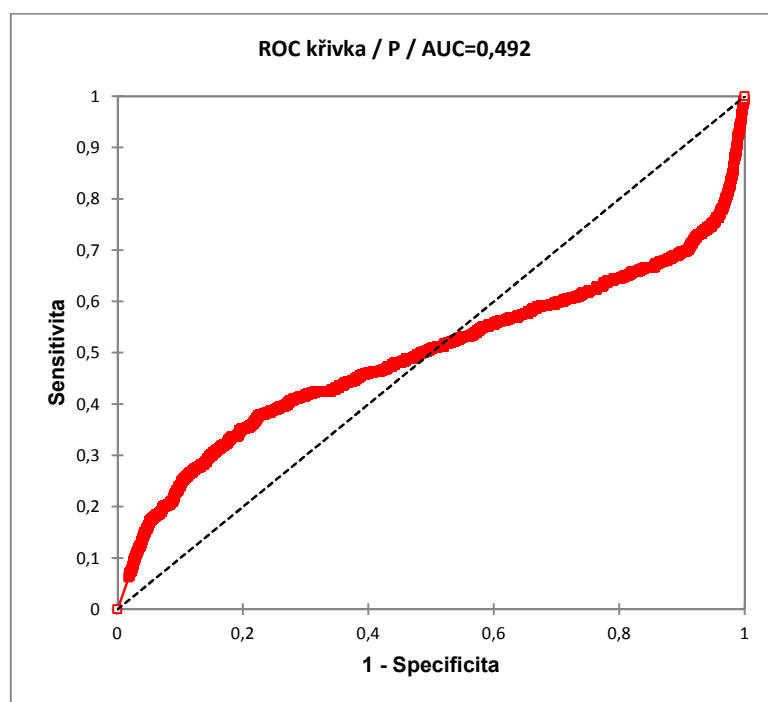
V tomto upraveném vzorku se vyskytuje celkem 67 460 aktivních společností a 680 společností bankrotních. Pro zjištění úspěšnosti predikce vytvořených modelů, je nejdříve nutné aplikovat rovnice č. 13 a 14 na nová data. Po vypočtení hodnot  $Q$  pro každou společnost, jsou tyto hodnoty postupně vkládány do pravděpodobnostní rovnice (rovnice č. 3). Následně jsou data podrobeny ROC analýze a je nutno zjistit nejen jak vypadá výsledná křivka ROC a hodnota AUC, ale i jaká úspěšnost predikce se váže k dříve určeným hladinám zlomu. Tento proces je proveden nejen pro model EU\_3.2, ale i pro model EU\_2.2.

Jako první je verifikace provedena pomocí modelu EU\_3.2, jehož graf je vyobrazen na obrázku č. 15. Tato křivka popisuje situaci, kdy zpočátku model predikuje slibně, ale v pozdější fázi je každý bod senzitivity vykoupen vyšším zvýšením falešné positivity. Tento vývoj samozřejmě není ideální, o čemž vypovídá i hodnota koeficientu AUC ve výši 0,492. V této chvíli je predikční schopnost modelu de facto na úrovni náhodného hádání. Pro účely pozdějšího srovnání mezi verifikacemi jednotlivých modelů, jsou uvedeny i výsledky úspěšnosti predikce jednotlivých kategorií pro hladinu zlomu určenou jak součtem senzitivity a specifity, tak pomocí nevyšší přesnosti.

Tab. 39 Úspěšnost verifikace 1-3

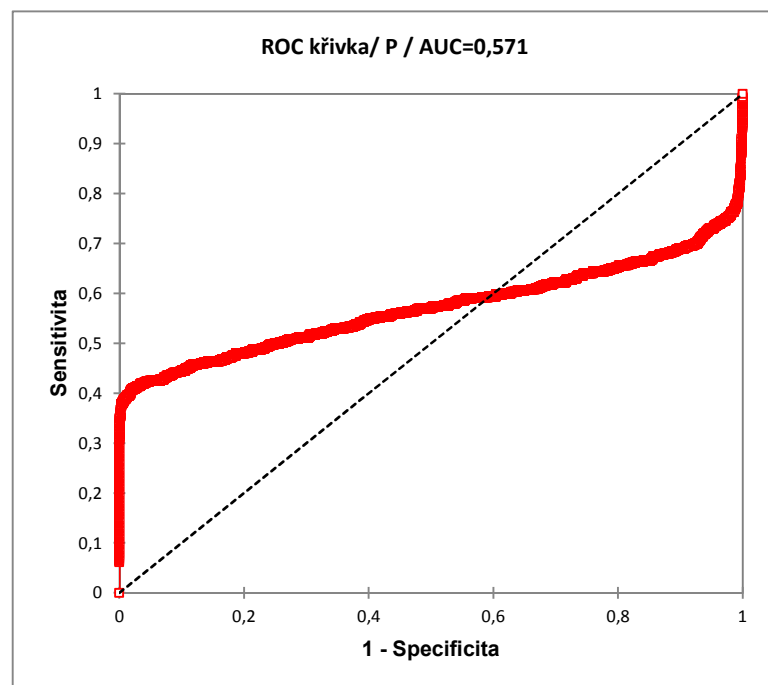
P=0,108		P=0,498	
Celková	75,30 %	Celková	82,50 %
Zdravé	75,66 %	Zdravé	82,97 %
Bankrotní	38,38 %	Bankrotní	31,91 %

Opět se potvrzuje myšlenka, že určení úspěšnosti na základě hladiny zlomu v bodě nejvyšší senzitivity a specifity vede k mnohem vyrovnanějším výsledkům než na základě celkové přesnosti.



Obr. 15 ROC: Verifikace 1-3

Dále je tento model verifikován na upraveném data setu, kde stejně jako při tvorbě těchto modelů, byla upravená zdravá složka data setu a byly vyfiltrovány společnosti s pochybnými finančními výkazy.



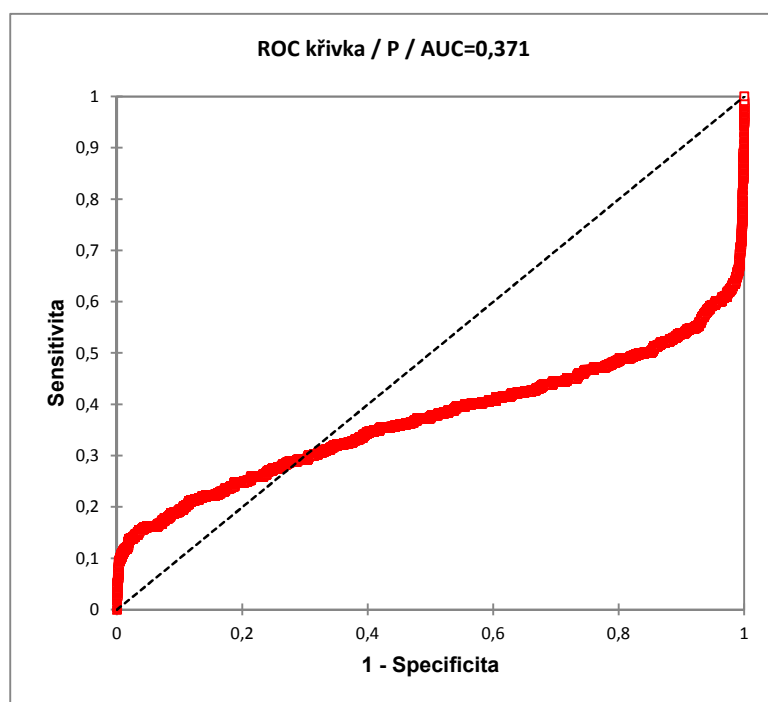
Obr. 16 ROC: Verifikace 2-3

Na této křivce je vidět, že vyfiltrováním společností s pochybnými výkazy, bylo za-  
mezeno růstu falešné positivity (označení zdravé společnosti za bankrotní) v první  
části grafu, což vede k lepšímu sklonu křivky a hodnoty AUC. V pozdější části se už  
znovu projevuje stejná tendence jako u původních dat, kdy zvýšení senzitivity vede  
k vyššímu zvýšení falešné positivity. Tento jev, kdy křivka ROC protíná diagonálu  
směrem dolů, je možné zdůvodnit výrazným nepoměrem zdravé a bankrotní slož-  
ky data setu nebo jejich podobností. Pokud jsou si tyto dvě složky příliš podobné (v  
hodnotách proměnných), tak pro odhalení další bankrotní společnosti je nutné  
zpřísnit kritéria pro rozřazení (hladina zlomu), jehož důsledkem je ale označení  
zdravých společností s podobnými daty za bankrotní (zvýšení falešné positivity).  
Tato situace napovídá tomu, že účetní proměnné možná nedokáží uspokojivě po-  
psat situaci bankrotujícího podniku, což bude zohledněno v závěru práce.

Tab. 40 Úspěšnost verifikace 2-3

P=0,108		P=0,498	
Celková	97,20 %	Celková	97,60 %
Zdravé	99,31 %	Zdravé	100,00 %
Bankrotní	38,38 %	Bankrotní	32,06 %

Tato tabulka nepřináší žádné závratné změny, pouze znovu potvrzuje fakt, že po  
očistění data setu od špatných zdravých společností došlo k prudkému nárůstu  
úspěšnosti predikce v této kategorii.



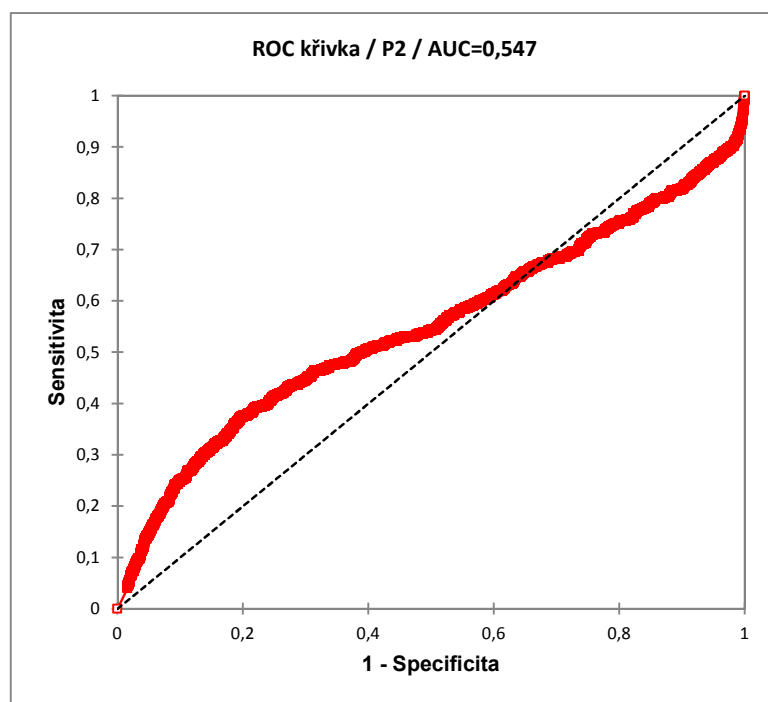
Obr. 17 ROC: Verifikace 3-3

Dále je nutné verifikovat i poslední možnou úpravu data setu na které byl tvořen právě používaný model EU\_3.2. Na obr. č. 17 je vidět, že odstraněním bankrotních společností, která neporušila ani jednu kritickou hladinu stanovenou pro zdravé společnosti (podle výkazů vypadají jako zdravé) se velmi zhoršila celková kvalita modelu. Toto zjištění je možno zdůvodnit tím, že v daném data setu existuje mnoho společností, které mají jednu z vybraných proměnných za kritickou hladinou a zbylé na úrovni zdravých společností. V tomto případě by model nebyl schopný rozlišit bankrotní společnosti od zdravých a současná úprava data setu by byla značným krokem vedle od cesty k nevhodnějšímu modelu. Tento fakt lze pozorovat i na snížených hodnotách AUC a úspěšnosti predikce bankrotních společností (následující tabulka).

Tab. 41 Úspěšnost verifikace 3-3

P=0,108		P=0,498	
Celková	97,30 %	Celková	97,70 %
Zdravé	99,32 %	Zdravé	100,00 %
Bankrotní	10,16 %	Bankrotní	0,23 %

Pokud by byl výše zmíněný předpoklad pravdivý, tak by logicky model vytvořený na druhé úpravě data setu (model EU\_2.2) měl dosahovat lepších výsledků při verifikaci na nových datech. Proto je předchozí postup u modelu EU\_3.2 následně proveden i pro model EU\_2.2.



Obr. 18 ROC: Verifikace 1-2

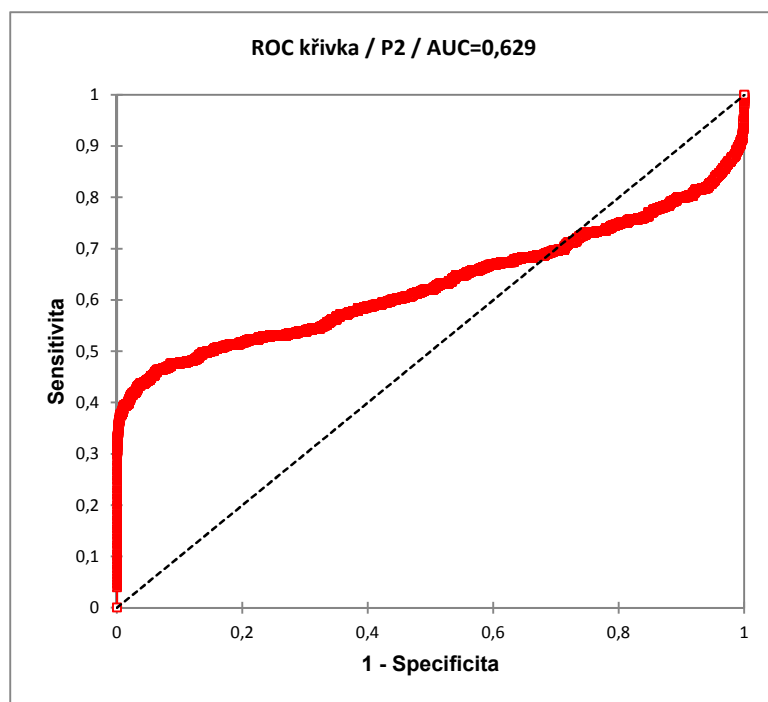
Obr. č. 18 už popisuje počátek verifikace modelu EU\_3.2, kdy je tento model aplikován na nová neupravená data. Už při zhodnocení grafického tvaru křivky je jasné, že je vhodnější než u předchozího modelu, o čemž se dá lehce přesvědčit pomocí vyšší hodnoty AUC (0,547). Tato hodnota se sice stále pohybuje okolo hodnoty 0,5, ale podle kritérií uvedených v části metodika se tento model dá považovat za kvalitní. Jednotlivé úspěšnosti pro hladiny zlomu určených celkovou přesností a součtem senzitivity se specifivitou jsou uvedeny níže.

Tab. 42 Úspěšnost verifikace 1-2

	P=0,092		P=0,5
Celková	79,10 %	Celková	87,30 %
Zdravé	79,51 %	Zdravé	87,90 %
Bankrotní	37,65 %	Bankrotní	27,50 %

Při srovnání úspěšnosti mezi modelem EU\_3.2 a EU\_2.2 je vcelku jasné, že model EU\_2.2 je na původním neupraveném data setu úspěšnější. Tento trend je nutné otestovat i na jeho dalších úpravách.

Jako další krok je tedy tento model otestován na datech s upravenou zdravou částí.



Obr. 19 ROC: Verifikace 2-2

Tento graf popisuje zatím nejkvalitnější verzi aplikování vytvořeného modelu na verifikační data set. Hodnota 0,629 u koeficientu AUC je značně vzdálená od

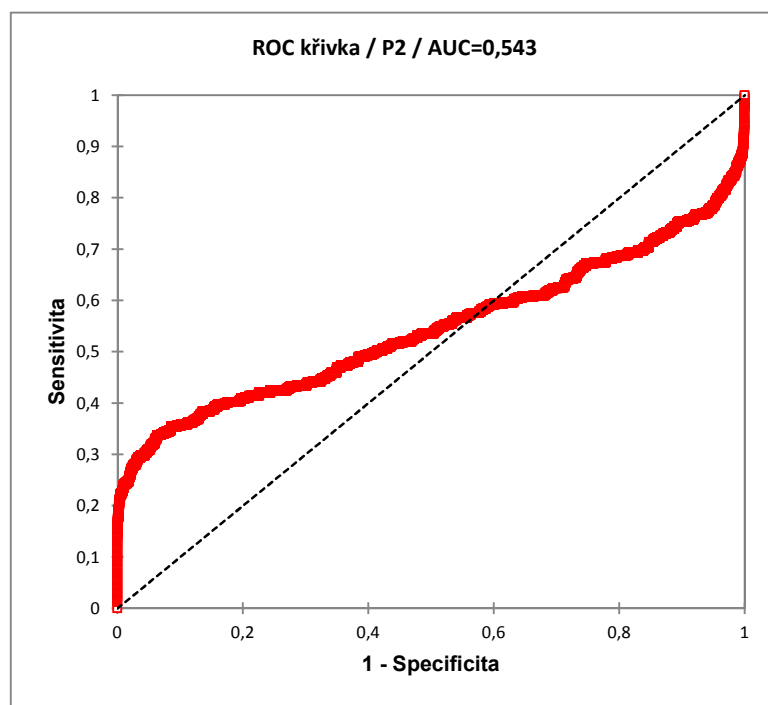
úspěšnosti náhodného hádání a dá se tedy považovat za poměrně úspěšný kvalitní model.

Tab. 43 Úspěšnost verifikace 2-2

	P=0,092		P=0,5
Celková	97,10 %	Celková	97,40 %
Zdravé	99,33 %	Zdravé	100,00 %
Bankrotní	37,65 %	Bankrotní	27,35 %

Matoucí jsou ovšem výsledky úspěšnosti pro dříve určené hladiny zlomu. I když je tento model mnohem kvalitnější než model EU\_3.2, tak hodnoty úspěšnosti predikce jsou téměř porovnatelné. Tento fakt vede k domněnce, že pro tento data set je původní model kvalitní, ale hladina zlomu u nových dat má odlišné optimum. Vzhledem k tomu, že tento jev nijak neovlivní úspěšnost zkoumané verifikace, tak nebude rozveden v této části, ale v části porovnání vytvořeného modelu a modelů od známých autorů, právě na těchto datech.

Pro jistotu je provedena verifikace i na data setu s upravenou bankrotní složku, i když se zde čekají horší výsledky.



Obr. 20 ROC: Verifikace 3-2

I když v této úpravě došlo ke zhoršení kvality modelu, tak se pořád dá označit za kvalitní, jelikož jeho hodnota AUC je ve výši 0,543. Toto zhoršení lze pozorovat hlavně na zhoršení predikce u bankrotních společností ilustrovaného v následující tabulce.

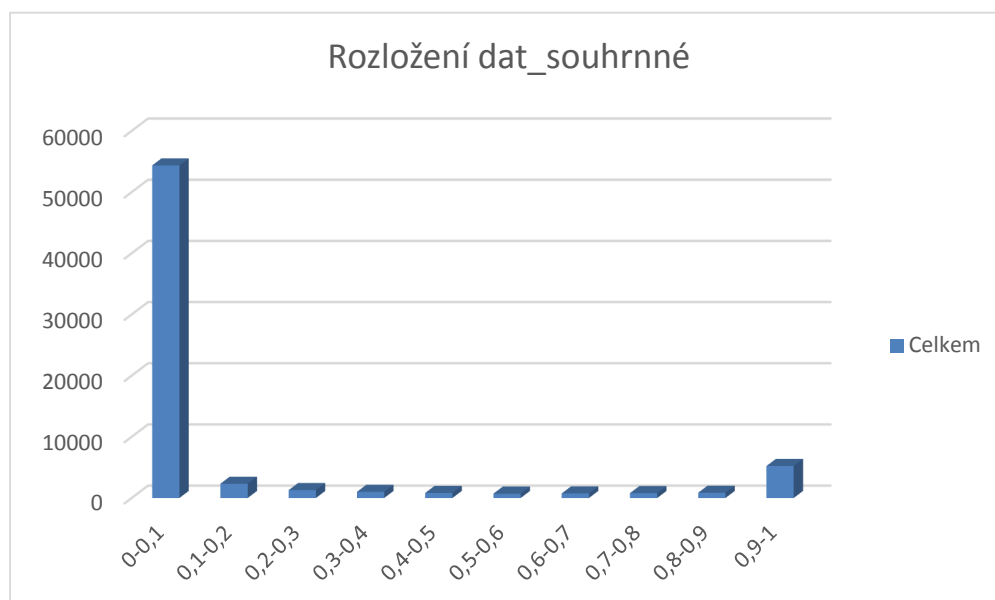
Tab. 44 Úspěšnost verifikace 3-2

P=0,092		P=0,5	
Celková	97,60 %	Celková	98,00 %
Zdravé	99,33 %	Zdravé	100,00 %
Bankrotní	22,40 %	Bankrotní	11,78 %

Závěrem této verifikace je tedy skutečnost, že i když je model EU\_3.2 částečně kvalitní (aplikace na data set s upravenou zdravou složkou), tak je téměř na úrovni náhodného hádání a o jeho kvalitě se dá spekulovat. Na druhou stranu model EU\_2.2 se projevil jako kvalitní ve všech případech, přičemž v data setu s upravenou zdravou složkou je nejkvalitnější. Z těchto výsledků lze vyvodit že třetí úprava dat, která se zabývala vymazáním bankrotních společností, které měly údaje porovnatelné s firmami zdravými, je zbytečná.

Jako nejuspěšnější se projevil model EU\_2.2 v první úpravě data setu s hladinou zlomu při  $P=0,092$  (určena na základě senzitivity a specifity). Model je kvalitní i na neupravených datech, ale pro zkvalitnění výsledků je vhodné provést úpravu zdravé složky data setu, ke které se při pozdějším testování bude přistupovat se zvýšenou pozorností.

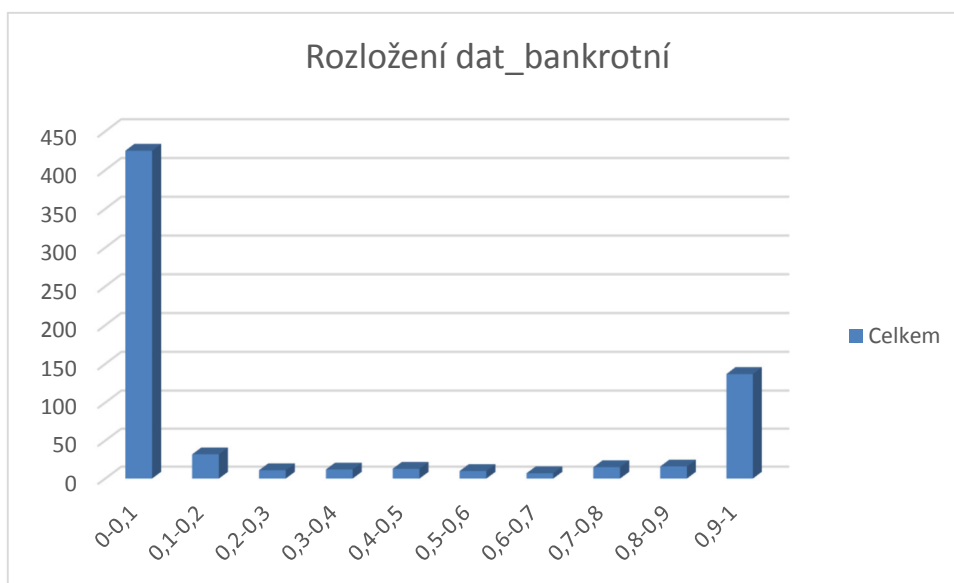
Vzhledem k tomu, že během testování modelu EU\_2.2 byly zjištěny vysoké rozdíly mezi počty společností spadajících do jednotlivých hladin zlomu, tak je na místě, aby byl vytvořen sloupcový graf, na kterém tyto rozdíly budou viditelné. Jelikož v pozdějším srovnání modelů bude pracováno pouze s neupravenými daty, tak i následující graf je vytvořen pro data, po aplikování modelu EU\_2.2, na zmíněný neupravený data set.



Obr. 21 Rozložení dat: EU\_1-2\_souhrnné



Na uvedeném sloupcovém grafu je vidět, že drtivá většina společností se nachází v intervalu  $<0; 0,1)$  a také na druhé straně v intervalu  $<0,9; 1)$ . Tento jev není nijak zvláštní, pokud je bráno v úvahu, že bankrotní složka je malá a může spadat právě do posledního intervalu, přičemž zdravá složka by byla obsažena v prvním a jev bankrotu by tak byl jednoznačně rozlišitelný. Tato myšlenka je ale neslučitelná s výsledky úspěšnosti predikce bankrotních společností, a proto je vytvořen nový graf, který už vyobrazuje pouze bankrotující společnosti.



Obr. 22 Rozložení dat: EU\_1-2\_bankrotní

Pokud by platila předchozí myšlenka, tak tento graf by obsahoval největší počet společností v posledních sloupcích, ale z grafu je viditelné, že skutečnost je odlišná. Nejvíce společností je v nejnižších hladinách pravděpodobnosti. Již dříve byly v textu zmíněny myšlenky o tom, že bankrotní společnosti v tomto data setu mají nezvykle dobré údaje ve finančních výkazech a mnohdy jsou na úrovni zdravých společností. Výše uvedený graf tuto myšlenku silně podporuje, ale pokud by tomu tak opravdu bylo, tak by příčina bankrotu u většiny zbankrotovaných společností byla jiná než špatný finanční stav společnosti, který by se propal do finančních výkazů.

Tento výstup poukazuje na možnost, že účetní predikční modely nemusí být tak významné ve spojitosti k predikci bankrotu, jak se doposud myslelo. Tato příčina by se mohla hledat v jiných datech, jako jsou makroekonomické údaje daného územního celku, historický vývoj společnosti aj. Navíc s ohledem na značný nepochopení v zastoupeních zdravých a bankrotních společností, by bylo patrně vhodným dalším rozšířením zvážit pro tyto případy reformulaci ROC algoritmu, jako např. Swensson et al. (2001), avšak s ohledem na cíl práce, nebyl věnován prostor pro detailní rozbor specifik modifikovaných odhadů ROC křivek. Vzhledem k tomu, že byl testován finální model pouze na jednom data setu, který shromažďoval data pouze z jednoho roku, tak se nedá tento výstup považovat za prokázaný a bylo by

potřeba jej otestovat na dalších datech, aby se potvrdil. Tato myšlenka bude zasažena do kontextu a více rozvinuta v závěru práce.

## 5.4 Srovnání se známými modely

Tato podkapitola se celá zabývá aplikací známých modelů na vybraná data a jejich srovnání s finálním modelem, jenž byl vytvořen v této práci (model EU\_2.2). K tomuto srovnání je ale nutné vybrat vhodný data set, na kterém se dají výsledky objektivně porovnat.

Vzhledem k tomu, že dat sety z ČR disponují velmi malou bankrotní složkou a na původním evropském data setu byl vytvářen finální model, tak je možné modely porovnávat pouze na novém evropském data setu. Tímto novým evropským data setem může být již získaný verifikační data set, ke kterému už existují i výsledky finálního modelu, a proto je třeba pouze aplikovat v praxi využívané modely a následně je porovnat.

Díky tomu, že finální model vytvořený v této práci sdílí většinu proměnných s Ohlsonovým modelem, tak je téměř nutností s tímto modelem a jeho modifikací vytvořit srovnání. Na druhou stranu Altmanův model je použit pouze jako doplňkové srovnání, ať už díky nízké podobnosti obou modelů nebo díky důvodům, které jsou uvedeny v části Aplikace: Altman.

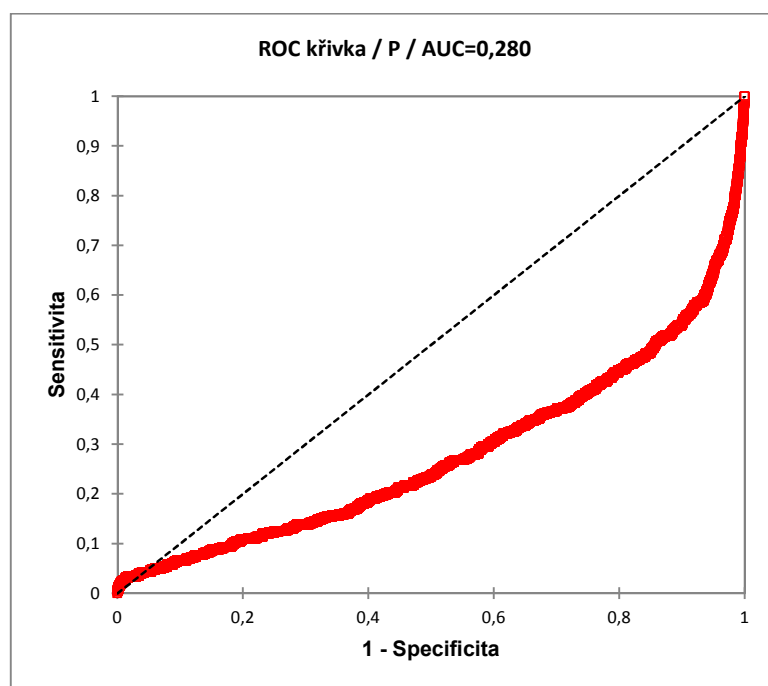
Faktem zůstává, že autoři Altman a Ohlson při tvorbě svých predikčních modelů nebrali v úvahu úpravu původních dat, která byla postupně zkoušena u modelu EU i CZ, a proto je srovnání provedeno pouze na neupravených datech verifikačního data setu, kde by měly modely O-score a Z-score dosahovat nejvyšších kvalit.

### 5.4.1 Aplikace: Ohlson

Pro aplikaci na vybraném data setu, je vybrán původní model z roku 1980, ale vzhledem k tomu, že se neočekává jeho vysoká predikční schopnost, tak je přidán i novější model z roku 2003. Oba tyto modely byly zkonstruovány samotným autorem na datech amerických společností. Existují sice i novější verze tohoto modelu, ale ty nebyly vytvořeny samotným autorem a jsou přizpůsobeny přímo jednotlivým prostředím států, jako je Turecko, Írán atd.

Předpoklad toho, že původní Ohlsonův model bude mít horší predikční schopnost, navazuje na často zmiňovanou vlastnost účetních modelů, a to horšící se predikční schopnost s větším časovým odstupem mezi aplikovaným modelem a daty, na které se model uplatňuje. Tento všeobecně uznávaný předpoklad lze lehce ověřit porovnáním kvality původního a modifikovaného modelu.

Obrázek č. 23 zobrazuje aplikaci ROC analýzy na výsledná data při použití původního Ohlsonova modelu z roku 1980. Z tvaru křivky a hodnoty AUC (0,280) je možno tvrdit, že tento model je nekvalitní, hluboko pod krajní hodnotou AUC=0,5.



Obr. 23 ROC: Ohlson\_1980

Samotné úspěšnosti jednotlivých kategorií jsou uvedeny v následující tabulce.

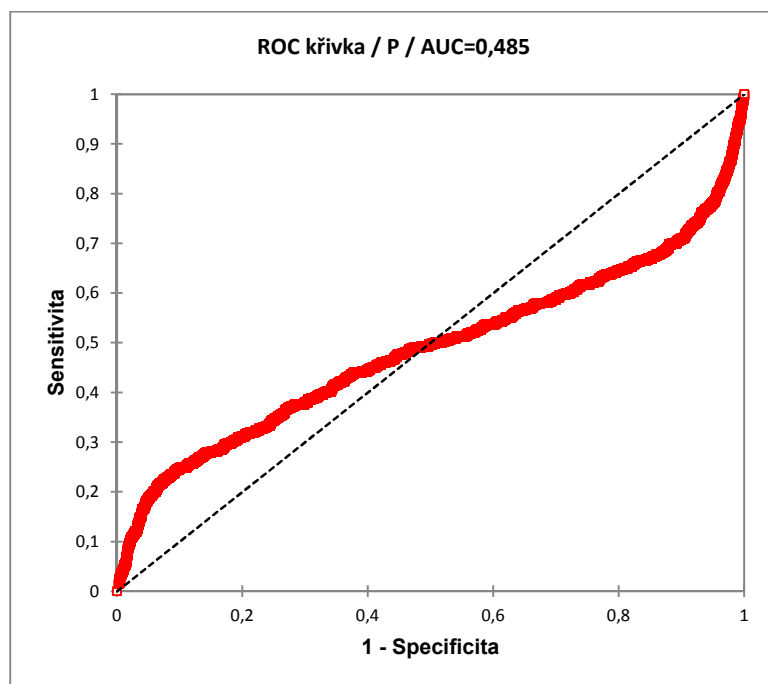
Tab. 45 Úspěšnost predikce Ohlson 1980

P=0,5		P=0,944	
Celková	58,10 %	Celková	97,60 %
Zdravé	58,47 %	Zdravé	98,53 %
Bankrotní	19,41 %	Bankrotní	3,24 %

Pro původní model Ohlson určil hladinu zlomu na hodnotu 0,5, ale pomocí ROC analýzy bylo zjištěno, že nejvýhodnější hladina, z hlediska součtu senzitivity a specifity, je v bodě 0,944. I když je značný rozdíl mezi oběma hladinami, tak je jasné, že ani jedna varianta není lepší, než finální model EU\_2.2.

S horším výsledkem se u tohoto modelu počítalo, a proto je nutné porovnat finální model hlavně s novější modifikací Ohlsonova modelu z roku 2003. ROC analýza výsledných dat je vyobrazena v grafu č. 24.

Z grafu je čitelné značné zlepšení oproti původnímu modelu z roku 1980. Tento novější model téměř dosahuje hladiny 0,5 koeficientu AUC a je viditelný i podobný trend křivky, jako u finálního modelu EU\_2.2.



Obr. 24 ROC: Ohlson\_2003

Výsledné úspěšnosti predikce pro jednotlivé kategorie jsou uvedeny v tabulce.

Tab. 46 Úspěšnost predikce Ohlson 2003

	P=0,5		P=0,734
Celková	65,50 %	Celková	89,80 %
Zdravé	65,80 %	Zdravé	90,50 %
Bankrotní	40,44 %	Bankrotní	24,41 %

V této tabulce jsou již značně lepší hodnoty a úspěšnost predikce u jednotlivých hladin se velmi podobá výsledkům u modelu EU\_2.2. Jednotlivé odlišnosti a zhodnocení úspěšnosti mezi modely je rozepsána až v části porovnání modelů.

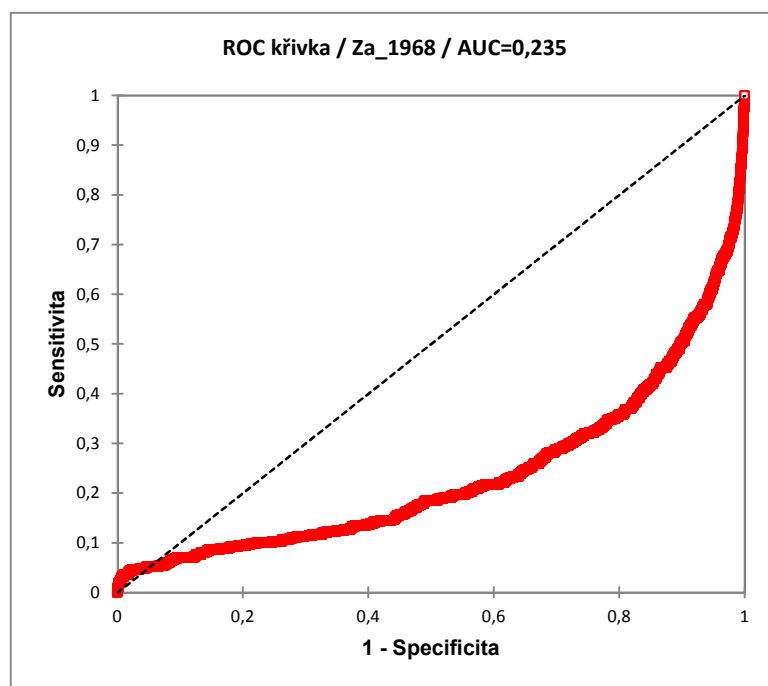
#### 5.4.2 Aplikace: Altman

Pro kvalitní porovnání je nutno aplikovat i model vytvořený Altmanem, ale nejdříve je nutné zmínit několik důležitých faktů. Původní modely Z-score braly v úvahu nerozdělený zisk, který v databázi Amadeus není možno získat. Proto je tato položka nahrazena ziskem čistým a o srovnání takto upraveného modelu se nedá uvažovat jako o srovnání finálního modelu z této práce a modelu Z-score. Toto srovnání je spíše obecného charakteru, přičemž je sledována pouze schopnost odlišných proměnných vysvětlit předložená data. Jinými slovy se jedná o snahu zjistit, jestli některá z proměnných uvedených v těchto modelech nezajistí kvalitnější popis dat, než je u vytvořeného predikčního modelu. Proto je naprosto nesmyslné uvádět a později porovnávat úspěšnost a součet senzitivity se specificitou, když

k tomuto zjištění stačí pouhá ROC analýza. U tohoto typu modelů proto bude zhodnocen pouze tvar křivky v grafu ROC a v závislosti na něm i hodnota AUC.

Pro porovnání byl tedy vybrán upravený model z roku 1968 a jeho nejnovější aktualizace samotným autorem v upravené formě z roku 1995. Je nutné zmínit, že sám Altman vytvořil nový model i pro rok 2010, ale tento model je modelem ratingovým nikoliv bankrotním, a proto nebude brán v úvahu.

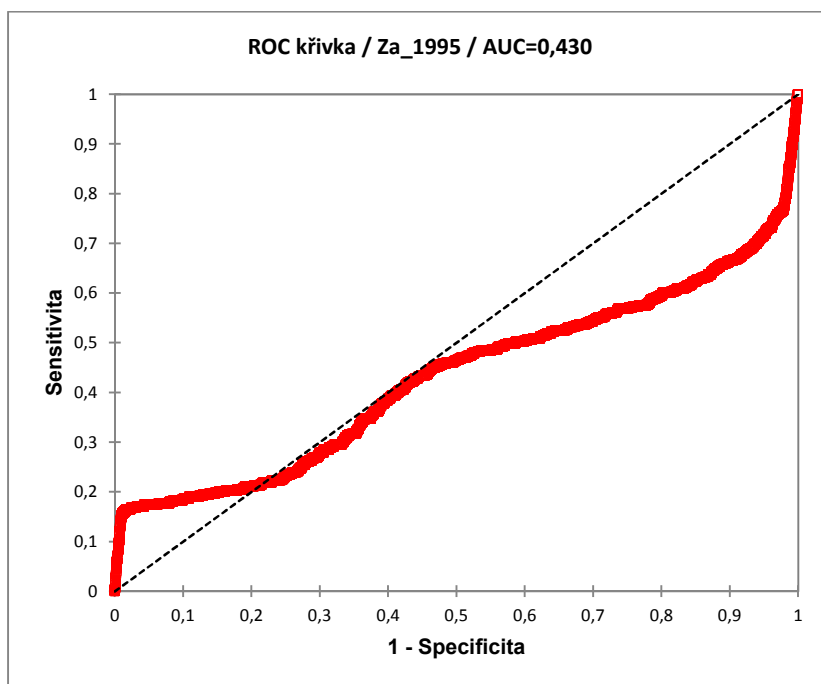
Aplikace ROC analýzy na výsledky modelu z roku 1968 je v následujícím grafu:



Obr. 25 ROC: Altman\_1968

Tento graf vypovídá o podobné situaci jako u původního Ohlsonova modelu, tedy o nekvalitním modelu, jenž není v současnosti vhodný k predikci bankrotu. Důležité pro pozdější porovnání je pouze zjištění hodnoty AUC ve výšce 0,235. Graf č. 26 na následující straně, už vykresluje ROC analýzu pro novější aktualizaci tohoto modelu.

Novější verze Altmanova modelu je opět mnohem lepší oproti verzi původní, ale i tak nedosahuje požadovaných hodnot pro kvalitní model. S mírným odchýlením je zde opět možné zpozorovat stejný tvar křivky jako u modelu EU\_3.2 a Ohlsonovy aktualizace z roku 2003. Hodnota AUC ve výšce 0,43 se pouze přiblížila hraniční hodnotě 0,5 a opět dosahuje nižší hodnoty než finální model. Dá se tedy spekulovat o tom, že žádná z proměnných uvedených v těchto porovnávaných modelech nevysvětluje bankrot natolik, aby bylo nutné uvažovat o jejím zařazení do vytvořeného modelu EU\_2.2, za účelem zlepšení jeho predikční schopnosti.



Obr. 26 ROC: Altman\_1995

### 5.4.3 Porovnání modelů

Po aplikaci modelů O-score a upraveného Z-score je nutné komplexně zhodnotit jejich kvalitu a predikční schopnost oproti predikčnímu modelu s pracovním označením EU\_2.2, který byl vytvořen v této práci. Nejdůležitější údaje potřebné pro komplexní porovnání zmíněných modelů jsou uvedeny v následující tabulce, kde zkratky Ú\_x vyjadřují úspěšnost predikce celkovou, zdravých a bankrotních společností.

Tab. 47 Porovnání modelů při aplikaci na verifikační data set

Název	AUC	P	Ú_c	Ú_z	Ú_b	S+S
EU_2.2	<b>0,547</b>	0,092	79,1 %	79,5 %	37,7 %	<b>1,17</b>
		0,107	87,3 %	87,9 %	27,5 %	<b>1,18</b>
O_1980	0,28	0,5	58,1 %	58,5 %	19,4 %	0,78
		0,944	97,6 %	98,5 %	3,2 %	1,02
O_2003	0,485	0,5	65,5 %	65,8 %	40,4 %	1,06
		0,734	89,8 %	90,5 %	24,4 %	1,15
A_1968	0,235	-	-	-	-	-
A_1995	0,43	-	-	-	-	-

V této tabulce jsou uvedeny nejen veškeré srovnávané modely a jejich údaje, ale jsou zde i u každého aplikovaného modelu porovnány původní a ideální hladiny zlomu. Původní hladina zlomu je udávána autory na původním data setu, přičemž

ideální hladina zlomu je jen pro srovnání a udává, jaká je pro tento data set nejvhodnější hladina, ve které se maximalizuje součet senzitivity a specifity. Tento údaj bere v úvahu rozdílné velikosti bankrotní a zdravé složky data setu, a proto je brán jako spolehlivější ukazatel úspěšnosti predikce než jednotlivé kategorie úspěšnosti.

Tato nová ideální hladina byla vytvořena i pro model EU\_2.2, ale oproti rozdílnosti u ostatních modelů, zde se jednalo pouze o posun necelé setiny procenta u hladiny zlomu a o jednu setinu jednotky v součtu senzitivity a specifity. Tyto výsledky jsou téměř stejné a při porovnání s modely o-score, kde se jedná o posuny několika desetin procent, tak jsou tyto změny téměř zanedbatelné.

Z uvedených dat je viditelná i další závislost, a to rozdíl v původní a ideální hladině zlomu (a k nim patřícímu součtu S+S) u starších a novějších modelů. U staršího modelu O\_1980 je tento posun vcelku drastický, a i když je u jeho novější verze pořad výrazný, tak je mnohem nižší. U „nejnovější verze“, za kterou by se dal považovat model EU\_2.2, je tento rozdíl téměř nepodstatný. Tato skutečnost opět potvrzuje teorii o tom, že starší modely jsou méně přesné a potřebují větší úpravu, aby dosahovaly lepších výsledků.

Z této tabulky se dá také jednoznačně potvrdit, že na základě porovnání koeficientů AUC je model EU\_2.2 kvalitnější a na základě součtu senzitivity a specifity přesnější v predikci bankrotu než srovnávané modely. Byla tedy získána odpověď na druhou výzkumnou otázku, kterou je, že vytvořený model je lepší v oblasti predikce bankrotu než zvolené známé predikční bankrotní modely.

## 5.5 Shrnutí výsledků

Při snaze o srovnání známých modelů a jejich modifikací bylo vhodné, aby byla pro daný účel vytvořena i nová modifikace unikátní pro tuto práci. Zmíněná modifikace byla zaměřena na středně velké společnosti ze stavebnictví v ČR, které nejsou obchodovány na akciových trzích.

Takto vzniklý model pro prostředí společností v České republice dosahuje velmi vysoké kvality, měřené na základě koeficientu AUC, spadajícího pod analýzu ROC. Jeho finální podoba je uvedena v rovnici č. 12 a dosahuje celkové úspěšnosti predikce až 96,41 % při hladině zlomu  $P=0,1$ . Tato zvolená hladina vyrovnává rozdíly mezi velikostí jednotlivých složek data setu (bankrotní a zdravá složka), a proto jsou i predikční schopnosti pro jednotlivé kategorie velmi vyrovnané. Predikční schopnost pro bankrotní společnosti dosahuje 96,43 % a pro zdravé 96,41 %.

Mnoho autorů se v odborné literatuře shoduje na názoru, že takto vytvořené predikční modely není možné posuzovat pouze na základě úspěšnosti predikce u původních dat, ale je nutné je verifikovat na datech nových. Proto byla provedena verifikace i tohoto modelu, ale data získaná pro nový data set byla nedostatečná. Při testování tohoto modelu bylo dosaženo téměř dokonalé schopnosti klasifikace, což je prakticky nemožný jev, který byl způsoben malým počtem společností v bankrotní složce nového data setu. V této situaci bylo možné pouze rozšířit původní omezení při tvorbě modelu, které se skládalo hlavně z omezení na jeden rok

a zemi. Při prozkoumání dat z jiných let se zjistilo, že v rozmezí pěti let zpětně je nedostatečný počet bankrotních společností, a proto by časové rozšíření nedávalo smysl. Jedinou možností, jak vytvořit predikční model na dostatečných datech, je v současné době pouze rozšíření o data z jiných zemí. Tato možnost byla zvolena a model byl znovu vytvořen pro evropské země, které spadají do EU nebo s ní mají velmi úzké vztahy.

Tento nový evropský model byl vytvořen stejným postupem, jako model český, jen se pracovalo s mnohanásobně vyšším počtem dat. Počet zkoumaných společností se v jednotlivých verzích modelu lišil, ale většinou se pohyboval v řádu desetitisíců společností. U tohoto modelu bylo uvažováno o dvou tvarech finální rovnice, ale nakonec byla pro finální model vybrána rovnice č. 13. Jeho celková úspěšnost při hladině zlomu  $P=0,092$  je 97,1 %, přičemž úspěšnost predikce zdravé složky dosahuje 97,44 % a složky bankrotní 91,1 %. Tento evropský model byl, stejně jako model český, označen za kvalitní skrze velmi vysokou hodnotu AUC.

Následně byla provedena verifikace na nových evropských datech a úspěšnost predikce i kvalita modelu výrazně klesla. Tento fakt je důvodem, proč jsou modely verifikovány, tudíž s tímto poklesem bylo počítáno. Verifikovaný model dosahuje sice nižších kvalit než model původní, ale pořád se dá podle hodnoty AUC zařadit do modelů kvalitních. Celková úspěšnost predikce, při aplikaci na nová data při hladině zlomu  $P=0,092$  je 97,1 %. Úspěšnost predikce u zdravé složky zůstala téměř totožná, a to 99,3 %. Naopak vysoký propad zaznamenala úspěšnost predikce bankrotní části. Tato úspěšnost dosahuje hodnoty pouhých 37,65 %, ale při bližším pohledu na zpracovávaná data, je jasné proč. Mnoho bankrotních společností v tomto vzorku dat dosahují stejných finančních výsledků, jako společnosti zdravé a je tedy možno spekulovat o tom, co je příčinou jejich bankrotu. Z tohoto důvodu není možné posuzovat úspěšnost daného modelu pouze na základě úspěšnosti predikce, ale je nutné jej porovnat s ostatními dnes využívanými modely.

Toto porovnání bylo provedeno hlavně s původní a modifikovanou verzí Ohlsonova bankrotního modelu, přičemž původní verze pochází z roku 1980 a modifikovaná z roku 2003. Altmanovy predikční modely nebylo možno celkově porovnat, protože v databázi Amadeus, ze které byly získány veškeré údaje, se neuvádí nerozdělený zisk, který je pro Altmanovi modely velmi důležitý. Tento údaj byl nahrazen ziskem čistým, ale Altmanův původní (1968) a modifikovaný (1995) model byl díky změně údajů porovnáván pouze z hlediska celkové kvality. Výsledkem tohoto porovnání je fakt, že původní modely zmíněných autorů nejsou dnes vůbec kvalitní a jejich modifikace se sice blíží kvalitním modelům, ale jsou spíše na úrovni náhodného hádání. Pro predikci bankrotu na tomto verifikačním evropském data setu, je tedy vytvořený model EU\_2.2 nejvhodnější variantou. Detailní vyobrazení tohoto srovnání se nachází na konci předchozí podkapitoly v tabulce č. 47. Pro úplnost této práce je vhodné porovnat i složení vytvořených modelů pro české i evropské prostředí s používanými bankrotními modely a jejich modifikacemi. Toto srovnání je uvedeno v tabulce č. 48 na následující straně. V závorkách pod hodnotami koeficientů u modelu CZ a EU jsou uvedeny p-hodnoty jednotlivých proměnných.



Tab. 48 Porovnání složení vytvořených a používaných modelů

	CZ	EU	O_1980	O_2003	A_1968	A_1995
Konstanta	-72,8 (0,009)	-24,8 (0,000)	-1,3	x	x	x
Čistý pracovní kapitál / Aktiva	64,1 (0,011)	9,9 (0,000)	-1,4	-1,25	1,2	6,56
EBIT / Aktiva	-109,2 (0,002)	x	x	x	3,3	6,72
Účetní hodnota VK / Závazky	3,7 (0,006)	1,8 (0,000)	x	x	0,6*	1,05*
Tržby / Aktiva	x	x	x	x	0,999	x
Log (Aktiva / deflátor HDP)	x	0,7 (0,000)	-0,4	-0,71	x	x
Závazky / Aktiva	x	16,3 (0,000)	6,02	2,2	x	x
KCZ / Oběžná aktiva	81,5 (0,008)	11,9 (0,000)	0,1	0,46	x	x
Závazky < Aktiva	x	x	-2,4	0,55	x	x
Čistý zisk / Aktiva	x	-12,4 (0,000)	-1,8	-3,79	1,4*	3,26*
(Čistý zisk + odpisy) / Závazky	x	x	0,3	-4,59	x	x
$(NI_t + NI_{t-1}) < 0$	x	-1,4 (0,000)	-1,7	0,157	x	x
$(NI_t + NI_{t-1}) / ( NI_t  +  NI_{t-1} )$	x	-2,6 (0,000)	-0,5	0,309	x	x

\*použité proměnné se neshodují s původní verzí v uvedeném modelu

I přes některé potíže při tvorbě nové modifikace bankrotního modelu a jejího následného porovnání, bylo dosaženo obou odpovědí na výzkumné otázky. Vytvořená modifikace bankrotního modelu má vyšší predikční sílu než náhodné hádání a je vhodnější z hlediska predikce bankrotu než jeho dnes používané vybrané alternativy.

## 5.6 Omezení a doporučení

Při tvorbě výše zmíněných modelů bylo několikrát nutné přijmout určité omezující předpoklady, aby bylo možné při modelování postupovat kupředu.

První pomyslné omezení nastalo v momentě, kdy byla vybrána databáze Amadeus, jako zdroj dat pro všechny data sety v této práci. Zvolením odlišné profesionální databáze by bylo možné získat odlišné údaje, které by mohly model silně ovlivnit. Navíc v této databázi není uváděna položka nerozdělený zisk, což značně

změnilo původní a modifikovaný Altmanův model, se kterým byl finální model této práce porovnáván. Za chybějící položku nerozděleného zisku byl totiž nahrazen zisk čistý, který nemá charakter akumulace a není tedy možné po porovnání tvrdit, že by vytvořený model EU\_2.2 měl vyšší vypovídací schopnost, než modely z-score. V této databázi bylo také odhaleno mnoho chybějících údajů a diskutabilně rozlišených firem (na zdravé a bankrotní), proto by bylo vhodné vytvořit identický model na odlišných datech z jiné databáze a finální modely porovnat.

Dalším omezením, co je nutné na závěr této práce zmínit, je použitelnost modelů vytvořených na čistě českých společnostech, tedy všechny modely s předponou CZ. Tyto modely jsou sice kvalitní a na původních datech úspěšné, ale tím že verifikující data set byl nedostačující, tak tyto modely nebyly nikdy úspěšně verifikovány. Zde se nachází prostor pro další práce, které by se mohly zabývat zajištěním dat pro vytvoření a verifikování kvalitního modelu, přímo pro prostředí českých společností.

Posledním omezením, které zde bude zmíněno, je využití finálního modelu EU\_2.2. Je pravda, že tento model je kvalitní a úspěšný nejen na původních, ale i na datech nových, což jej mimo jiné odlišuje od modelů z českého prostředí. Na druhou stranu, jeho verifikace byla provedena pouze na data setu z jednoho následujícího období. Bylo by tedy vhodné, aby se jeho úspěšnost potvrdila i na dalších data setech. Bohužel v době, kdy byla tvořena tato práce, byl data set z roku 2014 posledním robustním data setem, jenž byl použitelný pro verifikaci.

Vzhledem k pochybnostem o využitelnosti účetních proměnných pro predikci bankrotu, které byly v práci několikrát zmíněny, by bylo vhodné, aby se v budoucích výzkumech akademici zaměřili na alternativní obsah (tržní, makroekonomické a specifické proměnné) a formu (probit, BSM, neuronové sítě) těchto modelů a porovnali jejich úspěšnost s modely účetními.

## 6 Závěr

Cílem této závěrečné práce bylo porovnání známých bankrotních modelů a jejich modernějších modifikací. Pokud by šlo jen o porovnání známých existujících modelů, tak by tato práce měla minimální praktický přínos, a proto se zaměřuje i na tvorbu nové modifikace bankrotního modelu, která by kvalitně predikovala bankrot společností pro prostředí České republiky. Pro lepší přehlednost této práce, byla nejdříve vytvořena nová funkční modifikace modelu, která byla posléze zařazena do skupiny modelů určených pro porovnání.

Pro srovnání a jako podklad pro následnou modifikaci byly vybrány bankrotní modely od autor Jamese Ohlsona a Edwarda I. Altmana. Tyto modely se řadí do kategorie účetních modelů, a proto i vytvořená modifikace má účetní charakter. Jako obsah pro tento model byly vybrány účetní proměnné, které jsou obsaženy i v původních modelech obou výše zmíněných autorů. Tvar modelu byl odvozen od původního tvaru Ohlsonova modelu z roku 1980, jímž je logit model. Tento tvar je výhodný v tom, že jeho výstupem je přímo pravděpodobnost bankrotu testované společnosti.

Aby bylo možné tento nový model vytvořit, tak bylo nutné získat značné množství dat z mnoha společností, které by vytvořily potřebný data set. Tato data byla získána z databáze Amadeus od společnosti Bureau van Dijk a řadí se do dvou základních částí, kterými jsou data od bankrotních a zdravých společností. Vytvářený model byl zaměřen na středně veliké společnosti ze stavebního odvětví, které nejsou veřejně obchodovány. Po získání potřebných dat, bylo potřeba dopočítat hodnoty jednotlivých vysvětlujících proměnných, které jsou základem pro nový model. Tento model byl tvořen v programu Gretl a jeho průběžné hodnocení na základě ROC analýzy bylo prováděno v doplňku programu Excel s názvem XLSTAT. Vzhledem k faktu, že mnozí autoři ve světové literatuře doporučují vzniklé modely verifikovat na novém data setu, tak bylo nutné získat dodatečná data i pro tuto verifikaci.

Tímto způsobem byl vytvořen bankrotní model pro prostředí České republiky, který dosahoval velmi dobrých výsledků predikce. Ale už při jeho tvorbě byl zaznamenán fakt, že bankrotní složka původního data setu je nepoměrně malá oproti složce zdravé a výsledky modelu by tak nemuseli být obecně použitelné. Pro tyto účely se provádí otestování modelu na verifikačním data setu, ale Česká Republika v této oblasti v posledních několika letech nedisponuje tolika bankrotními společnostmi, aby byl vytvořen kvalitní data set s dostatečnou bankrotní složkou. Proto bylo nutné upustit od původních omezujících předpokladů a data set byl rozšířen na společnosti z celé Evropské unie.

Výsledkem byl model, který stejně jako model pro české společnosti dosahoval velmi dobré úspěšnosti při predikci bankrotu zkoumaných společností. Vzhledem k tomu, že u evropského data setu nehrozil nedostatek bankrotních společností, tak bylo model možné i kvalitně verifikovat. Podle očekávání úspěšnost predikce poklesla, ale i tak je model nadále kvalitní a dosahuje slušných výsledků.

V této fázi se otevřela možnost pro porovnání této vytvořené finální modifikace se známými používanými bankrotními modely. Tato skupina známých modelů a jejich modifikací byla otestována na verifikačním data setu pro evropský model a posléze byla porovnána s úspěšností vytvořeného evropského modelu. Výsledky tohoto porovnání jsou uvedeny v tabulce č. 47. v podkapitole srovnání modelů a vyplývá z nich, že vytvořený evropský model unikátní pro tuto práci, je kvalitnější a úspěšnější v predikci bankrotu než vybrané známé bankrotní modely. Pro srovnání je vhodné uvést, že nový vytvořený model dokázal správně klasifikovat 37,7 % bankrotních firem (při 79,1 % zdravých), přičemž původní Ohlsonův model pouze 19,4 %. Moderní verze Ohlsonova modelu dokázala správně klasifikovat dokonce 40,4 % bankrotních společností, ale to jen na úkor snížené přesnosti klasifikace zdravých společností (pouhých 65,8 %). Pokud by navíc na vytvořený model byla aplikována navrhovaná úprava zdravé složky dat, tak dosahuje téměř dokonalé klasifikace zdravých společností (99,3 %).

Vytvořením a porovnáním evropského modelu bylo dosaženo i obou žádaných odpovědí na výzkumné otázky této práce. Pro úplnost této práce bylo vhodné uvést i srovnání složení těchto zmíněných modelů, které je vyobrazeno v tabulce č. 48 v podkapitole shrnutí výsledků.

I když jsou tyto výsledky velmi dobré, tak nelze opomenout mnohá omezení, která byla v průběhu práce přijata. Tato omezení jsou souhrnně uvedena na konci předchozí kapitoly a dají se obecně rozdělit na tři druhy. Omezení formy a obsahu modelu, kde je rozebrána možnost odlišných a přípustně i lepších výsledků než za použití formy logit a účetních proměnných, o jejichž použitelnosti bylo v práci mnohokrát polemizováno. Dalším druhem je omezení vstupních dat, které bylo v této práci reprezentováno použitím databáze Amadeus. Je totiž pravděpodobné, že při využití jiných databází mohlo být dosaženo rozdílných výsledků a úspěšností. Obzvláště v návaznosti na značné množství pochybných údajů, které byly při práci se zmíněnou databází objeveny. Posledním pomyslným druhem omezení je omezení praktické využitelnosti vytvořených modelů, které se váže k nedostatečnému otestování modelů v praxi. Je pravda, že evropský model byl alespoň verifikován, ale i tato verifikace byla provedena pouze na datech z jednoho následujícího období.

Závěrem této práce je tedy úspěšné porovnání známých bankrotních modelů nejen s jejich novějšími modifikacemi, ale i s novou vytvořenou modifikací, která se prokázala jako nejvhodnější varianta pro predikci bankrotu na vybraných datech. Na druhou stranu tato práce poukazuje na mnohé nedostatky spojené s účetními bankrotními modely. I když se zde opět prokázalo, že účetní modely mají určitou predikční schopnost, tak je zde značný prostor pro zlepšení, který by se mohl zaplnit zvolením jiného typu modelu.

## 7 Literatura

- ADAMEC, J. *Faktory predikce finanční tísně*. In: sborník příspěvků mezinárodní vědecké doktorské konference INPROFORUM Junior, České Budějovice: Jihočeská univerzita, 2010, 5-9.
- ALTMAN, E. I. *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. The journal of finance, 1968, 23.4: 589-609.
- ALTMAN, E. I., HALDEMAN, R. G., NARAYANAN, P. *Zeta analysis. a new model to identify bankruptcy risk of corporations*. Journal of Banking and Finance, 1977, pp. 29-54. ISSN 0378-4266.
- ALTMAN, E. ET AL., *The Z-Metrics Methodology For Estimating Company Credit Ratings and Default Risk Probabilities*. New York: RiskMetrics Group., 2010.
- BEAVER, WILLIAM H. *Financial Ratios As Predictors of Failure*. Journal of Accounting Research 4, 1966, 71-111. DOI:10.2307/2490171.
- BEAVER, W. H., *Alternative accounting measures as predictors of failure*. The Accounting Review, 1968, 43(1), 113-122.
- CHAVA, S., & JARROW, R., *Bankruptcy prediction with industry effects*. Review of Finance 8, 2004, 537-569.
- CROSBIE, P., & BOHN, J. *Modeling default risk*. KMV corporation, 2003.
- ČÁMSKÁ, D., J. HÁJEK, *Companies related to the glass making industry and their financial health*. Transakční náklady českých ekonomických subjektů v insolvenčním řízení, možnosti jejich snižování a zlepšení statistiky insolvenčních řízení, 2012, 11.
- DLUHOŠOVÁ, D., ZMEŠKAL, Z., *Z-Metrics methodology and Altman rating models evolution*. Ostrava: 8 th International scientific conference Financial management of firms and financial institutions, 2011.
- FAWCETT, TOM. *An introduction to ROC analysis*. Pattern recognition letters, 2006, 27.8: 861-874.
- FITZPATRICK, P. J., *Symptoms of industrial failures*. Catholic University of America Press, 1931.
- GRICE, JOHN STEPHEN, INGRAM, ROBERT W., *Tests of the Generalizability of Altman's Bankruptcy Prediction Model*, Journal of Business Research, 2001, Volume 54, p.53.
- GRICE, JOHN STEPHEN A MICHAEL T. DUGAN. *Re-estimation of the Zmijewski and Ohlson Bankruptcy Prediction Models*. Advances in accounting, 2003, Volume 20, s. 77-93. ISBN 0762310669.
- HEIJ, C. *Econometric methods with applications in business and economics*. New York: Oxford University Press, 2004. ISBN 0199268010.
- HILLEGEIST, S., KEATING, E., CRAM, D., & LUNDSTEDT, K., *Assessing the probability of bankruptcy*. Review of Accounting Studies, 2004, 9, 5-34.

- JIN, M., *A Re-examination of Two Major Bankruptcy Prediction Models*. Vancouver: The University of British Columbia. Department of Commerce, 1993.
- KLEČKA, J., SCHOLLEOVÁ, H., *Bankruptcy models enuntiation for Czech glassmaking firms*. Economics and Management, 2010, pp. 954-959. ISSN 1822-6515
- KORDLAR, A. E. AND N. NIKBAKHT, *Comparing Bankruptcy Prediction Models in Iran*. School of Doctoral Studies (European Union) Journal, 2011, pp. 24-31.
- KUBÍČKOVÁ, DANA. *Ohlson's Model and its Prediction Ability in Comparison with Selected Bankruptcy Models in Conditions of Czech SMEs*. Economic Studies, 2015, 9(2), 155-173. ISSN 1802792X.
- KUPILÍK, O., *Analýza vlastnosti bankrotního modelu Ohlson*. Bakalářská práce, Plzeň Západočeská univerzita, Fakulta aplikovaných věd, 2013, 75 s.
- MCDONALD, R., *Derivatives Markets*. first edn., 2002. ISBN: 032128030X
- MERTON, R., *On the Pricing of Corporate Debt: The Risk Structure of Interest Rates*, Journal of Finance, 1974, 29, 449—70.
- MUZIR, E., CAGLAR, N., *The Accuracy of Financial Distress Prediction Models in Turkey: A Comparative Investigation with Simple Model Proposals*. 2009.
- NEUMAIEROVÁ, I. A NEUMAIER, I., *Index IN05: Index IN05*. In Evropské finanční systémy: Sborník příspěvků z mezinárodní vědecké konference. Brno: Masarykova univerzita v Brně, 2005. s. 143-146.
- OHLSON, J. A., *Financial ratios and the probabilistic prediction of bankruptcy*. Journal of Accounting Research, 1980, New York: 18(1), 109–131.
- PLATT, D. H., PLATT, M. B., *Development of a class of stable predictive variables: The case of bankruptcy prediction*. Journal of Business Finance & Accounting, 1990, 17(Spring), 31–51.
- SALIMI, A. Y., *Validity of Altmans Z-score model in predicting bankruptcy in recent years*. Academy of Accounting, 2015, 19(2), 233-238. ISSN 10963685.
- SHUMWAY, T., *Forecasting bankruptcy more accurately: a simple hazard model*. Journal of Business, 2001,74, 101–124.
- SCHOLLEOVÁ, H. *Ekonomické a finanční řízení pro neekonomy*. 1. vyd.: Grada Publishing, a. s., 2008. ISBN 978-80-247-2424-9.
- SWENSSON, R. G., KING, J. L., GUR, D., *A constrained formulation for the receiver operating characteristic (ROC) curve based on probability summation*. Medical physics, 2001, 28.8: 1597-1609.
- ŠLÉGR, P., *The Evaluation of Financial Stability of Czech Companies through the Z Score and the IN05 Index and their Comparison*. In: proceedings of the 7th WSEAS International Conference Recent Advances in Management, Marketing and Finances. Cambridge, MA USA, 2013, pp. 29-33.
- TAFFLER, RICHARD J. *The z-score approach to measuring company solvency*. The Accountant's Magazine, 1983, 87.921: 91-96.

- VALECKÝ, J., SLIVKOVÁ, E., *Mikroekonomický scoringový model úpadku českých podniků*. Ekonomická Revue: Central European Review of Economic Issues, 2012, 15(1), 15-26. ISSN 12123951.
- VASSALOU, M., & XING, Y. *Default risk in equity returns*. Journal of Finance, 2004, 59, 831–868.
- WANG, Y., CAMPBELL M., *Financial Ratios and the Prediction of Bankruptcy: The Ohlson's model Applied to Chinese publicly traded companies*, the Journal of organizational leadership & business, 2010, 1-15.
- WU, Y., C. GAUNT, S. GRAY, *A comparison of alternative bankruptcy prediction models*. Journal of Contemporary Accounting & Economics, 2010, Vol. 6, pp. 34-45.
- XU, M., ZHANG C., *Bankruptcy Prediction: The Case of Japanese Listed Companies*. Review of Accounting Studies, 2009, 14(4), 534-558. ISSN: 13806653.
- ZMIJEWSKI, M. E., *Methodological issues related to the estimation of financial distress prediction models*. Journal of Accounting Research, 1984, 24, 59–82.





# **Přílohy**



## A Literární přehled

Tab. 49 Altmanovo Z-metrics (Dluhošová, 2011)

Stupeň	Rating	Kotované firmy		Nekotované firmy	
		1 rok	5 roků	1 rok	5 roků
Vysoký	ZA+	0,02%	0,75%	0,03%	1,0%
	ZA	0,04%	1,25%	0,05%	1,5%
	ZA-	0,06%	1,75%	0,08%	2,0%
	ZB+	0,09%	2,50%	0,13%	3,0%
	ZB	0,14%	3,50%	0,20%	4,0%
	ZB-	0,20%	4,50%	0,30%	5,0%
Střední	ZC+	0,30%	6,00%	0,45%	6,5%
	ZC	0,50%	9,00%	0,70%	9,0%
	ZC-	1,00%	14,00%	1,50%	13,0%
Nízký	ZD+	2,00%	20,00%	3,00%	20,0%
	ZD	4,00%	30,00%	5,00%	30,0%
	ZD-	10,00%	45,00%	10,00%	45,0%
	ZF+	25,00%	65,00%	18,00%	55,0%
	ZF	50,00%	80,00%	30,00%	65,0%
	ZF-	100,00%	100,00%	100,00%	100,0%

Tab. 50 Souhrnný přehled současných modifikací O-score

	<b>X<sub>1</sub></b>	<b>X<sub>2</sub></b>	<b>X<sub>3</sub></b>	<b>X<sub>4</sub></b>	<b>X<sub>5</sub></b>	<b>X<sub>6</sub></b>	<b>X<sub>7</sub></b>	<b>X<sub>8</sub></b>	<b>X<sub>9</sub></b>	<b>Konstanta</b>
<b><i>Koeficient</i></b>	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	$\beta_9$	$\beta_0$
<b>Ohlson (1980)</b>	-0,41	6,03	-1,43	0,08	-2,47	-1,83	0,29	-1,72	-0,52	-1,32
<b>Jin (1993)</b>	-0,17	1,75	-0,85	0,04	-0,29	-2,50	-2,36	0,95	-0,52	-2,25
<b>Jin (1993)</b>	-0,16	0,87	-2,06	-0,22	-0,09	-6,10	-1,66	-0,13	-0,36	-0,73
<b>Ohlson (2003) A</b>	-0,77	3,22	-0,32	0,59	0,04	-2,86	-2,85	0,37	0,21	0
<b>Ohlson (2003) B</b>	-0,88	3,93	0,05	0,17	0,65	-0,55	-2,89	0,66	-0,3	0
<b>Ohlson (2003) P</b>	-0,71	2,20	-1,25	0,46	0,55	-3,79	-4,59	0,16	0,31	0
<b>Muzira a Cağlara (2009)</b>	-0,23	7,19	-0,07	0,61	-1,71	3,26	-4,19	0,44	-0,15	-4,58
<b>Wu et al. (2010)</b>	-0,17	3,69	-1,87	0	-0,54	0,03	-0,06	1,16	-1,02	-7,2
<b>Wang et al. (2010) 1</b>	-0,89	0,95	-0,92	2,4e <sup>-3</sup>	2,95	-0,011	-0,03	3,21	0,59	-1,31
<b>Wang et al. (2010) 2</b>	-0,28	-0,21	-0,21	-0,021	1,47	-7,5e <sup>-3</sup>	-0,05	-4,16	-0,93	-5,52
<b>Wang et al. (2010) 3</b>	-0,59	-0,49	-0,47	-1,6e <sup>-3</sup>	2,00	-0,01	-0,04	3,72	-0,18	-2,48
<b>Wang et al. (2010) 1'</b>	0	-0,99	-0,99	2,4e <sup>-3</sup>	3,38	0	0	3,11	0	-7,51
<b>Wang et al. (2010) 2'</b>	0	-0,14	-0,16	-0,023	1,53	0	0	4,29	0	-7,43
<b>Wang et al. (2010) 3'</b>	0	-0,41	-0,41	1,7e <sup>-3</sup>	2,18	0	0	3,86	0	-6,77
<b>Kordlar a Nikbakht (2011)</b>	-0,14	14,58	2,92	-0,6	-0,17	-1,4	-2,6	3,79	-0,25	-12,87

## B Vyhledávací kritéria pro databázi Amadeus

Tab. 51 Amadeus EU 2013 výsledky

	Mezivýsledky	Výsledky
Status: Aktivní	19 304 742	19 304 742
NACE Rev. 2 hlavní sekce: F. Stavebnictví	2 488 426	2 098 446
Kategorie společností dle velikosti: Středně velké společnosti	2 789 348	290 059
Standardizovaná právní forma: Společnost s ručením omezeným, veřejná obchodní společnost, živnostník	20 148 639	256 131
Poslední účetní období: 2015, 2014, 2013	19 430 948	<b>235 774</b>
	<b>Mezivýsledky</b>	<b>Výsledky</b>
Status: Likvidace, Bankrot, Likvidace s rozpuštěním společnosti, Bankrot s rozpuštěním společnosti	793 824	793 824
NACE Rev. 2 hlavní sekce: F. Stavebnictví	2 488 426	115 502
Kategorie společností dle velikosti: Středně velké společnosti	2 789 348	15 712
Standardizovaná právní forma: Společnost s ručením omezeným, veřejná obchodní společnost, živnostník	20 148 639	14 049
Poslední účetní období: 2013	4 263 132	<b>3 749</b>

Tab. 52 Amadeus EU 2014 výsledky

	<b>Mezivýsledky</b>	<b>Výsledky</b>
Status: Aktivní	19 304 742	19 304 742
NACE Rev. 2 hlavní sekce: F. Stavebnictví	2 488 426	2 098 446
Kategorie společností dle velikosti: Středně velké společnosti	2 789 348	290 059
Standardizovaná právní forma: Společnost s ručením omezeným, veřejná obchodní společnost, živnostník	20 148 639	256 131
Poslední účetní období: 2015, 2014	15 167 816	<b>185 338</b>
	<b>Mezivýsledky</b>	<b>Výsledky</b>
Status: Likvidace, Bankrot, Likvidace s rozpuštěním společnosti, Bankrot s rozpuštěním společnosti	793,824	793,824
NACE Rev. 2 hlavní sekce: F. Stavebnictví	2,488,426	115,502
Kategorie společností dle velikosti: Středně velké společnosti	2,789,348	15,712
Standardizovaná právní forma: Společnost s ručením omezeným, veřejná obchodní společnost, živnostník	20,148,639	14,049
Poslední účetní období: 2014	13,902,036	<b>2,674</b>

## C Model CZ

### Model CZ\_1.1

Model 1: Logit, za použití pozorování 1-2292 (n = 2284)

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu					Předpovězené		
	koeficient	směr. chyba	z	p-hodnota		0	1
<b>const</b>	-4,241	1,935	-2,191	0,028	<b>Skutečné</b>	0	2256
<b>X1</b>	-1,662	1,024	-1,623	0,105		1	28
<b>X2</b>	-0,917	2,304	-0,398	0,690			
<b>X3</b>	-0,243	0,196	-1,240	0,215			
<b>X4</b>	-0,142	0,168	-0,844	0,399			
<b>X5</b>	0,528	0,512	1,031	0,303			
<b>X6</b>	-1,114	0,964	-1,155	0,248			
<b>X7</b>	-0,486	0,319	-1,522	0,128			
<b>X8</b>	1,816	0,587	3,095	0,002			
<b>X9</b>	-1,326	2,370	-0,560	0,576			
<b>X10</b>	0,602	0,600	1,003	0,316			
<b>X11</b>	0,202	1,469	0,137	0,891			
<b>X12</b>	-0,207	0,814	-0,254	0,800			
Střední hodnota závisle proměnné				0,012	Sm. odchylka závisle proměnné		0,110
McFaddenův koeficient determinace				0,192	Adjustovaný koeficient determinace		0,106
Logaritmus věrohodnosti				-122,047	Akaikovo kritérium		270,094
Schwarzovo kritérium				344,632	Hannan-Quinnovo kritérium		297,279

### Model CZ\_1.2

Model 1: Logit, za použití pozorování 1-2292 (n = 2284)

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu					Předpovězené		
	koeficient	směr. chyba	z	p-hodnota		0	1
<b>const</b>	-5,017	0,270	-18,569	0,000	<b>Skutečné</b>	0	2262
<b>X8</b>	2,275	0,398	5,714	0,000		1	28
<b>X9</b>	-0,900	0,313	-2,872	0,004			
Střední hodnota závisle proměnné				0,012	Sm. odchylka závisle proměnné		0,110
McFaddenův koeficient determinace				0,135	Adjustovaný koeficient determinace		0,115
Logaritmus věrohodnosti				-130,763	Akaikovo kritérium		267,526
Schwarzovo kritérium				284,737	Hannan-Quinnovo kritérium		273,802

**Model CZ\_2.2 Všechna omezení neupraveno**

Model 1: Logit, za použití pozorování 1-2292 (n = 2284)

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu					Předpovězené			
	koeficient	směr. chyba	z	p-hodnota				
						0	1	
<b>const</b>	-125,537	67,057	-1,872	0,061	<b>Skutečné</b>	0	474	0
<b>X1</b>	110,775	59,416	1,864	0,062		1	3	25
<b>X2</b>	-113,753	79,989	-1,422	0,155				
<b>X3</b>	3,308	2,601	1,272	0,204				
<b>X4</b>	-1,498	1,785	-0,839	0,401				
<b>X5</b>	1,547	3,016	0,513	0,608				
<b>X6</b>	13,702	12,656	1,083	0,279				
<b>X7</b>	121,655	67,630	1,799	0,072				
<b>X9</b>	-19,567	55,527	-0,352	0,725				
<b>X10</b>	16,684	17,471	0,955	0,340				
<b>X11</b>	2,518	10,098	0,249	0,803				
<b>X12</b>	1,098	5,932	0,185	0,853				
Střední hodnota závisle proměnné				0,056	Sm. odchylka závisle proměnné		0,230	
McFaddenův koeficient determinace				0,907	Adjustovaný koeficient determinace		0,796	
Logaritmus věrohodnosti				-10,017	Akaikovo kritérium		44,035	
Schwarzovo kritérium				94,658	Hannan-Quinnovo kritérium		63,896	

**Model CZ\_2.2**

Model 1: Logit, za použití pozorování 1-2292 (n = 2284)

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu					Předpovězené			
	koeficient	směr. chyba	z	p-hodnota				
						0	1	
<b>const</b>	-72,787	28,000	-2,600	0,009	<b>Skutečné</b>	0	474	0
<b>X1</b>	64,095	25,303	2,533	0,011		1	3	25
<b>X2</b>	-109,161	35,882	-3,042	0,002				
<b>X3</b>	3,704	1,366	2,712	0,007				
<b>X7</b>	81,450	30,497	2,671	0,008				
Střední hodnota závisle proměnné				0,056	Sm. odchylka závisle proměnné		0,230	
McFaddenův koeficient determinace				0,879	Adjustovaný koeficient determinace		0,832	
Logaritmus věrohodnosti				-13,109	Akaikovo kritérium		36,218	
Schwarzovo kritérium				57,311	Hannan-Quinnovo kritérium		44,493	



## D Model EU

### Model EU\_1.1

Model 1: Logit, za použití pozorování 1-91728 (n = 91083)

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu

	koeficient	směr. chyba	z	p-hodnota	Skutečné	Předpovězené	
						0	1
<b>const</b>	-4,488	0,195	-23,024	0,000	0	89817	6
<b>X1</b>	-0,130	0,052	-2,517	0,012	1	1257	3
<b>X2</b>	0,871	0,205	4,254	0,000			
<b>X3</b>	-0,003	0,003	-1,155	0,248			
<b>X4</b>	0,000	0,003	0,107	0,915			
<b>X5</b>	0,043	0,054	0,794	0,427			
<b>X6</b>	-0,090	0,058	-1,533	0,125			
<b>X7</b>	0,000	0,000	0,905	0,365			
<b>X8</b>	1,824	0,072	25,272	0,000			
<b>X9</b>	-0,922	0,202	-4,563	0,000			
<b>X10</b>	-0,008	0,005	-1,591	0,112			
<b>X11</b>	-0,142	0,223	-0,634	0,526			
<b>X12</b>	-0,714	0,124	-5,782	0,000			
Střední hodnota závisle proměnné				0,014	Sm. odchylka závisle proměnné		0,117
McFaddenův koeficient determinace				0,131	Adjustovaný koeficient determinace		0,129
Logaritmus věrohodnosti				-5775,073	Akaikovo kritérium		11576,146
Schwarzovo kritérium				11698,600	Hannan-Quinnovo kritérium		11613,465

### Model EU\_1.2

Model 1: Logit, za použití pozorování 1-91728 (n = 91087)

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu

	koeficient	směr. chyba	z	p-hodnota	Skutečné	Předpovězené	
						0	1
<b>const</b>	-4,537	0,038	-120,082	0,000	0	89820	7
<b>X1</b>	-0,061	0,021	-2,912	0,004	1	1258	2
<b>X2</b>	0,915	0,198	4,613	0,000			
<b>X8</b>	1,805	0,070	25,812	0,000			
<b>X9</b>	-0,911	0,197	-4,624	0,000			
<b>X12</b>	-0,649	0,038	-17,061	0,000			
Střední hodnota závisle proměnné				0,014	Sm. odchylka závisle proměnné		0,117
McFaddenův koeficient determinace				0,130	Adjustovaný koeficient determinace		0,129
Logaritmus věrohodnosti				-5782,595	Akaikovo kritérium		11577,190
Schwarzovo kritérium				11633,708	Hannan-Quinnovo kritérium		11594,414

**Model EU\_2.1**

Model 1: Logit, za použití pozorování 1-21863

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu

	koeficient	směr. chyba	z	p-hodnota	Skutečné	Předpovězené	
						0	1
const	-25,025	1,675	-14,944	0,000	0	20585	18
X1	9,990	1,565	6,383	0,000	1	226	1034
X2	0,801	1,483	0,540	0,589			
X3	1,781	0,147	12,091	0,000			
X4	0,016	0,052	0,313	0,754			
X5	0,708	0,184	3,843	0,000			
X6	16,279	0,766	21,247	0,000			
X7	11,967	1,496	7,997	0,000			
X9	-13,402	2,214	-6,053	0,000			
X10	0,074	0,625	0,119	0,905			
X11	-1,397	0,392	-3,562	0,000			
X12	-2,573	0,247	-10,414	0,000			
Střední hodnota závisle proměnné				0,058	Sm. odchylka závisle proměnné		0,233
McFaddenův koeficient determinace				0,762	Adjustovaný koeficient determinace		0,760
Logaritmus věrohodnosti				-1144,777	Akaikovo kritérium		2313,554
Schwarzovo kritérium				2409,465	Hannan-Quinnovo kritérium		2344,798

**Model EU\_2.2**

Model 1: Logit, za použití pozorování 1-21863

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu

	koeficient	směr. chyba	z	p-hodnota	Skutečné	Předpovězené	
						0	1
const	-24,848	1,634	-15,208	0,000	0	20585	18
X1	9,944	1,526	6,515	0,000	1	226	1034
X3	1,786	0,121	14,730	0,000			
X5	0,677	0,161	4,199	0,000			
X6	16,259	0,756	21,500	0,000			
X7	11,928	1,460	8,169	0,000			
X9	-12,348	1,151	-10,726	0,000			
X11	-1,394	0,391	-3,563	0,000			
X12	-2,566	0,246	-10,439	0,000			
Střední hodnota závisle proměnné				0,058	Sm. odchylka závisle proměnné		0,233
McFaddenův koeficient determinace				0,762	Adjustovaný koeficient determinace		0,761
Logaritmus věrohodnosti				-1144,977	Akaikovo kritérium		2307,953
Schwarzovo kritérium				2379,886	Hannan-Quinnovo kritérium		2331,386

**Model EU\_3.1**

Model 1: Logit, za použití pozorování 1-21803

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu

	koeficient	směr. chyba	z	p-hodnota	Skutečné	Předpovězené	
						0	1
<b>const</b>	-23,494	1,724	-13,628	0,000	0	20591	12
<b>X1</b>	3,955	1,534	2,579	0,010	1	128	1072
<b>X2</b>	-3,874	1,958	-1,979	0,048			
<b>X3</b>	3,510	0,265	13,235	0,000			
<b>X4</b>	0,021	0,069	0,301	0,763			
<b>X5</b>	0,623	0,243	2,567	0,010			
<b>X6</b>	21,544	1,116	19,301	0,000			
<b>X7</b>	8,044	1,418	5,672	0,000			
<b>X9</b>	-2,101	3,451	-0,609	0,543			
<b>X10</b>	-16,146	1,679	-9,617	0,000			
<b>X11</b>	-1,969	0,514	-3,829	0,000			
<b>X12</b>	-2,951	0,329	-8,974	0,000			
Střední hodnota závisle proměnné				0,055	Sm. odchylka závisle proměnné		0,228
McFaddenův koeficient determinace				0,858	Adjustovaný koeficient determinace		0,856
Logaritmus věrohodnosti				-658,685	Akaikovo kritérium		1341,370
Schwarzovo kritérium				1437,247	Hannan-Quinnovo kritérium		1372,607

**Model EU\_3.2**

Model 1: Logit, za použití pozorování 1-21803

Závisle proměnná: Status

Směrodatné chyby založené na Hessiánu

	koeficient	směr. chyba	z	p-hodnota	Skutečné	Předpovězené	
						0	1
<b>const</b>	-23,465	1,681	-13,958	0,000	0	20590	13
<b>X1</b>	3,876	1,524	2,544	0,011	1	127	1073
<b>X2</b>	-4,427	1,672	-2,648	0,008			
<b>X3</b>	3,588	0,227	15,815	0,000			
<b>X5</b>	0,587	0,212	2,768	0,006			
<b>X6</b>	21,721	1,070	20,305	0,000			
<b>X7</b>	7,984	1,407	5,675	0,000			
<b>X10</b>	-16,718	1,363	-12,268	0,000			
<b>X11</b>	-1,955	0,512	-3,820	0,000			
<b>X12</b>	-2,952	0,328	-9,002	0,000			
Střední hodnota závisle proměnné				0,055	Sm. odchylka závisle proměnné		0,228
McFaddenův koeficient determinace				0,858	Adjustovaný koeficient determinace		0,856
Logaritmus věrohodnosti				-658,901	Akaikovo kritérium		1337,802
Schwarzovo kritérium				1417,700	Hannan-Quinnovo kritérium		1363,833