

**Univerzita Palackého v Olomouci**

**Přírodovědecká fakulta**

**Katedra geoinformatiky**

**ČASOPROSTOROVÉ PREDIKČNÍ MODELOVÁNÍ  
PANDEMIE COVID-19 UŽITÍM UMĚLÉ  
NEURONOVÉ SÍTĚ**

**Bakalářská práce**

**Martin KUKRÁL**

**Vedoucí práce Mgr. Karel MACKŮ, Ph.D.**

**Olomouc 2023**

**Geoinformatika a kartografie**

## **ANOTACE**

Tématem bakalářské práce je problematika tvorby umělé neuronové sítě pro potřeby časoprostorového predikčního modelování pandemie COVID-19 v administrativních jednotkách České republiky. K tomu jsou použita volně dostupná epidemiologická data sekvenční povahy, jež jsou za účelem usnadnění učení sítě nejprve analyzována vybranými statistickými metodami a transformována do vhodné podoby před vstupem do modelu. Sama tvorba konkrétní architektury umělé neuronové sítě plnící stanovené cíle je rozebrána z teoretického i praktického hlediska. Závěrečným krokem je vícefázová validace přesnosti predikcí i samotného nejlepšího dosaženého modelu zvolenými metodami užitím stanovených kritérií a vizualizací.

Výsledkem je kromě rozboru dané problematiky finální model, stejně jako nástroj vytvořený autorem práce v programovacím jazyce Python, který je použit k realizaci celého výše nastíněného procesu. Práce by měla reprezentovat jednu z možných cest časoprostorového predikčního modelování epidemiologických dat, případně by mohla také sloužit jako inspirace pro tvorbu podobných umělých neuronových sítí pro potřeby predikce sekvenčních dat i v jiných tematických oblastech.

## **KLÍČOVÁ SLOVA**

umělé neuronové sítě; strojové učení; predikce časových řad; Python; COVID-19

Počet stran práce: 41

Počet příloh: 7 (z toho 6 elektronických a 1 volná)



## **ANNOTATION**

This bachelor thesis deals with the design of an artificial neural network for the needs of spatiotemporal predictive modelling of the COVID-19 pandemic in the administrative units of the Czech Republic. For this purpose, freely available epidemiological data of sequential nature are used, which are first analysed using selected statistical methods and transformed into a suitable form before entering the model in order to facilitate learning of the network. The actual design of a specific artificial neural network architecture fulfilling the stated objectives is discussed from both theoretical and practical perspectives. The final phase is the multi-stage validation of the accuracy of the predictions and of the best achieved model itself, attained by the chosen methods using selected criteria and visualizations.

The results are, in addition to the analysis of the issue, the final model, as well as the tool created by the author of the thesis in the Python programming language, which was used to implement the entire process outlined above. The work should represent one of the possible ways of spatiotemporal predictive modelling of epidemiological data, while it could also serve as an inspiration for the creation of similar artificial neural networks for the needs of sequential data prediction in other subject areas.

## **KEYWORDS**

artificial neural networks; machine learning; time series forecasting; Python; COVID-19

Number of pages: 41

Number of appendices: 7

**Prohlašuji, že**

- bakalářskou práci včetně příloh jsem vypracoval samostatně a uvedl jsem všechny použité podklady a literaturu,
- jsem si vědom, že na moji bakalářskou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,
- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, bakalářskou práci užívat (§ 35 odst. 3),
- souhlasím, že údaje o mé bakalářské práci budou zveřejněny ve Studijním informačním systému UP,
- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užití výsledky a výstupy mé bakalářské práce v rozsahu § 12 odst. 4 autorského zákona,
- použít výsledky a výstupy mé bakalářské práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).



Martin Kukrál

V Olomouci dne 27. 4. 2023

Děkuji mému vedoucímu práce Mgr. Karlovi Macků, Ph.D. za důvěru při sestavování netradičně zaměřené práce a za hodnotné podněty a připomínky při jejím dokončování. Dále děkuji RNDr. Janu Brusovi, Ph.D. za přístup na počítač 3D laboratoře katedry, čímž mi ušetřil nemálo hodin výpočetního času při testování různých variant modelu a při tvorbě toho finálního. Velký dík patří také všem osobám z mého nejbližšího okolí, bez jejichž podpory a trpělivosti bych práci dokončil jen stěží.

# UNIVERZITA PALACKÉHO V OLMOUCI

Přírodovědecká fakulta

Akademický rok: 2021/2022

## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Martin KUKRÁL**  
Osobní číslo: **R200209**  
Studijní program: **B0532A330011 Geoinformatika a kartografie**  
Téma práce: **Časoprostorové predikční modelování pandemie COVID-19 užitím umělé neuronové sítě**  
Zadávající katedra: **Katedra geoinformatiky**

### Zásady pro vypracování

Cílem práce je sestavit predikční model založený na neuronové síti pro potřeby časoprostorového predikčního modelování. Proces bude zaměřen na epidemiologické téma vybraných charakteristik pandemie COVID-19 v administrativních jednotkách ČR. Student nejprve vybere vhodná volně dostupná data, která budou mít dostatečně detailní prostorový aspekt. Vstupní data budou zanalyzována metodami exploratorní analýzy s cílem postihnout jejich vnitřních závislostí pro tvorbu robustnějšího modelu. Nad daty bude sestaven model neuronové sítě (např. rekurentní neuronové vrstvy, Bayesovské neuronové a další) pomocí kterého budou modelovány vybrané charakteristiky z dostupných dat tak, aby byla možná přímá validace výsledků. Práce by měla reprezentovat jednu z možných cest časoprostorového predikčního modelování epidemiologických dat, případně by mohla také sloužit jako inspirace pro tvorbu podobných umělých neuronových sítí pro potřeby predikce sekvenčních dat i v jiných tematických oblastech.

Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data) bude odevzdána v digitální podobě na CD nebo jiném digitálním médiu, text práce bude spolu s vybranými přílohami odevzdán ve dvou svazcích na sekretariát katedry. O práci student vytvoří webové stránky, které budou v souladu s pravidly dostupnými na stránkách katedry. Diplomová práce bude zpracována podle zásad dle Voženilka (2002) a také podle šablon dostupných na stránkách katedry. Na závěr práce připojí student jednostránkové resumé v anglickém jazyce. Jako shrnutí bakalářské práce bude vytvořen poster.

Rozsah pracovní zprávy: **max. 50 stran**  
Rozsah grafických prací: **dle potřeby**  
Forma zpracování bakalářské práce: **tištěná**

### Seznam doporučené literatury:

- Cliff, A. D., & Ord, J. K. (1975) Space-Time Modelling with an Application to Regional Forecasting. Transactions of the Institute of British Geographers, (64), 119. <https://doi.org/10.2307/621469>
- Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., & Dehmer, M. (2020). An Introductory Review of Deep Learning for Prediction Models With Big Data. Frontiers in Artificial Intelligence, 3. <https://doi.org/10.3389/frai.2020.00004>
- Nikparvar, B., Rahman, M. M., Hatami, F., & Thill, J.-C. (2021). Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network. Scientific Reports, 11(1), 21715. <https://doi.org/10.1038/s41598-021-01119-3>
- Niraula, P., Mateu, J., & Chaudhuri, S. (2022). A Bayesian machine learning approach for spatio-temporal prediction of COVID-19 cases. Stochastic Environmental Research and Risk Assessment. <https://doi.org/10.1007/s00477-021-02168-w>
- Kruse, R., Borgelt, C., Braune, C., Mostaghim, S., & Steinbrecher, M. (2016). Computational Intelligence. <https://doi.org/10.1007/978-1-4471-7296-3>
- Voženílek, V., 2002. Diplomové práce z geoinformatiky. Olomouc: Vydavatelství Univerzity Palackého. ISBN 80-244-0469-9

Vedoucí bakalářské práce: **Mgr. Karel Macků, Ph.D.**  
Katedra geoinformatiky

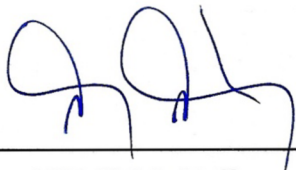
Datum zadání bakalářské práce: **5. května 2022**

Termín odevzdání bakalářské práce: **5. května 2023**

L.S.

---

**doc. RNDr. Martin Kubala, Ph.D.**  
děkan



---

**prof. RNDr. Vít Voženílek, CSc.**  
vedoucí katedry

V Olomouci dne 15. září 2022

# OBSAH

<b>SEZNAM POUŽITÝCH ZKRATEK .....</b>	<b>9</b>
<b>ÚVOD .....</b>	<b>10</b>
<b>1 CÍLE PRÁCE .....</b>	<b>11</b>
<b>2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY.....</b>	<b>12</b>
2.1 Přístupy k predikování COVID-19.....	12
2.2 Metodika Ministerstva zdravotnictví České republiky .....	14
2.3 Akademické práce v České republice .....	15
<b>3 METODY A POSTUP ZPRACOVÁNÍ .....</b>	<b>16</b>
3.1 Použité metody.....	16
3.2 Použitý software a hardware .....	17
3.3 Použitá data.....	17
3.4 Postup práce.....	18
<b>4 ANALÝZA A PŘEDZPRACOVÁNÍ VSTUPNÍCH DAT.....</b>	<b>19</b>
4.1 Stacionarita .....	19
4.2 Autokorelace .....	20
4.3 Datové transformace .....	21
<b>5 TVORBA UMĚLÉ NEURONOVÉ SÍTĚ.....</b>	<b>24</b>
5.1 Relevantní teoretické principy.....	24
5.1.1 Umělý neuron .....	25
5.1.2 Vrstva umělých neuronů .....	26
5.1.3 Umělá neuronová síť.....	26
5.1.4 Dropout .....	27
5.1.5 Rekurentní neuronové vrstvy .....	28
5.1.6 Bayesovská parametrizace vrstev.....	29
5.1.7 Ztrátové funkce.....	30
5.1.8 Algoritmus učení.....	32
5.1.9 Automatická optimalizace hyperparametrů .....	34
5.2 Třída B-GRU-PTSAR .....	35
<b>6 VALIDACE MODELU.....</b>	<b>39</b>
6.1 Validace automatické optimalizace hyperparametrů.....	39
6.2 Validace učení.....	41
6.3 Validace predikcí.....	42
<b>7 VÝSLEDKY.....</b>	<b>48</b>
<b>8 DISKUZE .....</b>	<b>49</b>
<b>9 ZÁVĚR.....</b>	<b>51</b>
<b>POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE</b>	
<b>PŘÍLOHY</b>	

## SEZNAM POUŽITÝCH ZKRATEK

<b>Zkratka</b>	<b>Význam</b>
AI	Artificial Intelligence
ANN	Artificial Neural Network
ADF	Augmented Dickey-Fuller
B-GRU-PTSAR	Bayesian Gated Recurrent Unit Parallel Time Series Autoregressor
BNN	Bayesian Neural Network
BP(TT)	Backpropagation (Through Time)
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CSV	Comma Separated Value
CUDA	Compute Unified Device Architecture
ČR	Česká republika
ELBO	Evidence Lower Bound
GIS	geografický informační systém
GPU	Graphics Processing Unit
GRU	Gated Recurrent Unit
KLD	Kullback-Leiblerova divergence
LISA	Local Indicators of Spatial Autocorrelation
LSTM	Long-Short Term Memory
MAE	Mean Average Error
MCMC	Markov Chain Monte Carlo
ML	Machine Learning
MSE	Mean Squared Error
MZČR	Ministerstvo zdravotnictví České republiky
OOP	objektově orientované programování
ORP	obec s rozšířenou působností
(P)ACF	(Partial) Autocorrelation Function
RNN	Recurrent Neural Network
(S)GD	(Stochastic) Gradient Descent
(S)ARIMA	(Seasonal) Autoregressive Integrated Moving Average
SLDB	Sčítání lidu, domů a bytů
S(E)IR	Susceptible, (Exposed), Infectious, Recovered
ÚZIS	Ústav zdravotnických informací a statistiky
VI	Variational Inference

# ÚVOD

O tvorbu predikčních modelů, jimiž je možné informovaněji předvídat vývoj nějaké veličiny, je velký zájem v mnoha různých oblastech, přičemž konkrétně v epidemiologii patří do popředí zájmu jak samotných epidemiologů, tak poskytovatelů zdravotní péče (Tabataba et al., 2017). Jednu z možných cest k jejich tvorbě představují metody strojového učení, konkrétně umělé neuronové sítě, které se v posledních letech dostávají do popředí světového výzkumu.

Motivací je snaha otestovat využitelnost těchto moderních technologií pro potřeby predikování časoprostorových epidemiologických dat, konkrétně na případu pandemie COVID-19 na území ČR (České republiky). Výsledný model sestavený tímto způsobem by měl být teoreticky schopen poskytovat velmi přesné predikce, z čehož by mohly potenciálně těžit relevantní české zdravotnické, výzkumné a administrativní instituce.

Práce v mnohém koncepčně navazuje na již realizovaný světový výzkum (viz podkapitola 2.1), nicméně v konkrétních aspektech je výsledkem autorova samostatného výzkumu a snahy navrhnout vhodné řešení.



# 1 CÍLE PRÁCE

Cílem bakalářské práce je sestavit predikční model založený na umělé neuronové síti pro potřeby časoprostorového predikčního modelování, konkrétně pro vybrané epidemiologické téma vybraných charakteristik pandemie COVID-19 v administrativních jednotkách ČR.

Nejprve budou vhodně vybrána volně dostupná data, která budou mít dostatečně detailní prostorový aspekt. Data budou nejprve před vstupem do umělé neuronové sítě analyzována a předzpracována statistickými metodami s cílem pochopit jejich vlastnosti k vytvoření robustnějšího modelu a usnadnění procesu učení sítě. Nad takto připravenými daty bude posléze sestaven co nejpřesnější model užitím vybraných relevantních teoretických metod umělých neuronových sítí tak, aby byla možná následná přímá validace výsledků i modelu samotného.

Práce by měla reprezentovat jednu z možných cest časoprostorového predikčního modelování epidemiologických dat, případně by mohla také sloužit jako inspirace pro tvorbu podobných umělých neuronových sítí pro potřeby predikce sekvenčních dat i v jiných tematických oblastech.

## 2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Spolu s dostupností rozsáhlých datových sad, hardwarových výpočetních prostředků, softwarových nástrojů a s rozvojem relevantní teorie zažívá oblast strojového učení (anglicky *machine learning*, dále ML) prudký rozvoj, v důsledku čehož jsou každoročně vyvíjeny nové modelové architektury a přístupy (Steinbach et al., 2022). To se přirozeně týká také problematiky tvorby predikčních modelů, v níž nejen umělé neuronové sítě (anglicky *artificial neural network*, dále ANN) prokázaly svou využitelnost v řadě odborných soutěží a představují významná zlepšení oproti tradičním metodám, především s ohledem na schopnost modelovat míru nejistoty (Bojer, 2022). Epidemiologie jakožto lékařský obor veskrze orientovaný na datovou analýzu čerpá z těchto nových poznatků při sestavování predikčních algoritmů pro šíření nemocí a modelování účinnosti léčby, díky čemuž je pokrok v oblasti ML vnímán jako produktivní úsilí s potenciálem na zlepšení zdraví populace (Wiemken a Kelly, 2020).

### 2.1 Přístupy k predikování COVID-19

Za historicky nejstarší a nejjednodušší metodu matematického modelování infekčních nemocí mohou být považovány tzv. kompartmentové modely, z nichž nejznámější SIR (*Susceptible, Infectious, Recovered*) a jeho deriváty jsou používány epidemiology dodnes, jelikož poskytují hodnotný prvotní vhled do chování nemoci (Rodrigues, 2016). Účinnost tohoto přístupu nicméně bývá častým cílem různých testování a srovnávání s modernějšími metodami, přičemž například Moein et al. (2021) ve své kazuistice zaujímají vůči SIR modelům spíše kritický postoj. Zároveň také existují snahy rozšířit dobře interpretovatelné SIR modely o nové poznatky z oblasti datového modelování, například Vega et al. (2022) popisují hybridní model užívající stochastický ML algoritmus pro aproximaci změn parametrů SIR modelu v čase. Za zmínku stojí, že různě modifikované SIR modely tvoří jádro metodiky MZČR (Ministerstvo zdravotnictví ČR) pro predikce vývoje pandemie COVID-19 na území ČR (viz podkapitola 2.2).

Využití nacházejí také obecně používané třídy lineárních autoregresních modelů časových řad známé z matematické statistiky, jmenovitě především ARIMA (*Autoregressive Integrated Moving Average*) nebo SARIMA (*Seasonal ARIMA*), je-li zahrnuta informace o sezónním průběhu dat. Kupříkladu Sahai et al. (2020) a Tan et al. (2022) využívají těchto metod k predikování trendu epidemiologických charakteristik COVID-19 vybraných států s relativně vysokou mírou přesnosti, z čehož plyne potenciál pro jejich využitelnost relevantními státními a zdravotními institucemi. Přestože ML metody jako ANN dosahují lepší přesnosti v predikcích než modely ARIMA, přináší s sebou tyto modely výrazně nižší celkovou komplexitu při tvorbě, výpočetní náročnost a také opadá potřeba jemného ladění architektury a hyperparametrů ANN, což z nich dělá technologii vhodnou pro jednodušší problémy (Zhou et al., 2020). Srovnání ARIMA a ANN v rámci predikčních modelů pro COVID-19 poskytují například Rguibi et al. (2022), kteří dosahují použitelných výsledků oběma metodami, nicméně ANN se prokázaly jako přesnější. Nelineární odraz principů ARIMA modelů čistě v rámci ANN lze nalézt u tzv. *nonlinear autoregressive neural networks*, pomocí nichž Saliaj a Nissi (2022) dosáhly významně vyšší přesnosti predikcí nově nakažených oproti lineárním modelům, jako byla mimo jiné právě ARIMA.

Sama o sobě velmi široká oblast ML algoritmů nabízí k řešení problematiky COVID-19 široké množství přístupů, jako jsou např. *support vector machines*, *random forests*, *k-nearest neighbours*, *gradient boosting* a samozřejmě celá podmnožina ANN se svými vlastními podoblastmi, z nichž za nejdůležitější z hlediska predikování lze považovat RNN

(*Recurrent Neural Network*), CNN (*Convolutional Neural Network*) a jejich další deriváty (Kamalov et al., 2022).

Přestože CNN dosáhly své popularity především v oblasti počítačového vidění a zpracování obrazu, je možné jimi v důsledku shodné počítačové reprezentace jakožto polí zpracovávat také časové řady (Wibawa et al., 2022). Tento přístup úspěšně aplikovali například Mohimont et al. (2021) a jejich výsledky byly následně integrovány do metodiky predikování COVID-19 ve francouzském regionu Grand-Est. Pokrok byl zaznamenán také ve tvorbě hybridních CNN-RNN predikčních modelů, přičemž kupříkladu Zain a Alturki (2021) efektivně využili výstupy konvolučních vrstev jako vstupy do rekurentní vrstvy, která z nich následně extrahovala sekvenční informace. Další velmi zajímavou práci využívající hybridní CNN-RNN model představili Muñoz-Organero a Queipo-Álvarez (2022) pro reprezentaci hodnot incidence uvnitř spádových oblastí zdravotnických zařízení madridského regionu jako spojitý povrch ve formě 2D matic (resp. rastrů), s nimiž následně pracovali podobně jako se snímky.

V rámci samotných RNN bývá velmi často užívána varianta LSTM (*Long-Short Term Memory*), jelikož řeší značné množství klíčových problémů základní, tzv. *vanilla* RNN architektury, díky čemuž jsou považovány za jednu z nevhodnějších možností pro tvorbu predikcí (Arora et al., 2020). Přímé srovnání těchto dvou architektur poskytují Alassafi et al. (2022) a konstatují, že užití LSTM vedlo ke zvýšení přesnosti predikcí případů COVID-19 o 5,13 %. Dále jsou prováděna komparativní testování různě modifikovaných LSTM, přičemž například Verma et al. (2022) dosáhli největší přesnosti predikcí užitím sítě s několika vzájemně napojenými, tzv. *stacked* LSTM vrstvami a výše zmíněným hybridním CNN-RNN, resp. CNN-LSTM modelem. Přesných predikcí s vysokým prostorovým detailem dosáhli také Lucas et al. (2022) užitím *stacked* LSTM sítě, do níž vstupovaly jak vybrané epidemiologické charakteristiky, tak informace o pohybu a chování obyvatel získané z dat ze sociální sítě Facebook. Dále Fritz et al. (2022) zvolili ve své koncepčně podobné práci cestu modelovat data o mobilitě obyvatel jakožto grafovou strukturu, nad níž byl následně sestaven hybridní model, v němž tzv. *graph neural network* část řešila prostorovou složku přesunu obyvatel a RNN část časovou. Hodnotné srovnání mezi LSTM a GRU (*Gated Recurrent Unit*), vývojově mladší alternativou pro LSTM, v kontextu COVID-19 poskytují ArunKumar et al. (2021) a konstatují kratší dobu učení a nižší výpočetní náročnost GRU vrstev, nicméně oba přístupy dosáhly velmi podobné míry úspěchu napříč několika stanovenými úkoly. Specifické využití GRU vrstev zvolili Bi et al. (2022) při tvorbě predikčního modelu, jehož výsledky byly následně zpracovány multipopulačním evolučním algoritmem pro nalezení optimálního anticovidového epidemiologického opatření za předpokladu maximalizace účinnosti a minimalizace nákladů.

S ohledem na komplexní časoprostorové dynamiky šíření nemoci COVID-19 není ani kvůli přítomnosti náhodného šumu v datech možné vytvářet perfektní modely, a proto je vhodné při jejich tvorbě určit korespondující míry nejistoty k predikcím (Cabras, 2021). Ačkoliv existuje mnoho studií zaměřených na predikování COVID-19 užitím ML metod, využití BNN (*Bayesian Neural Network*) je velmi omezené, přestože poskytují robustní metodologii pro kvantifikaci nejistoty při určování parametrů sítě, a tedy i predikcí (Chandra a He, 2021). Niraula et al. (2022) poskytují specifický hybridní model, v němž výstup LSTM vstupuje jako jedna složka do bayesovského časoprostorového regresního modelu pro predikování nových případů COVID-19, přičemž na vstupu byla použita epidemiologická, dopravní a demografická data. Dále Silk et al. (2022) využívají jednu z alternativ pro určování nejistoty predikcí, tzv. *ensemble* přístup, při němž jsou sofistikovanými metodami kombinovány výstupy z několika odlišných epidemiologických modelů.

Vhled do zajímavé problematiky automatizovaného ML v kontextu COVID-19 poskytují Abbasimehr a Paki (2020) využitím bayesovské optimalizace k nalezení optimálních hyperparametrů modelu.

Přehled náročných problémů během modelování pandemie COVID-19 shrnují Khan et al. (2022), z nichž za zmínku stojí např. nedostupnost dostatečného množství trénovacích dat, riziko náhodného šumu v datech, jejich mnohdy nestrukturovaná nebo zcela chybná evidence či vliv vládních opatření na šíření nemoci. Další komplikace při predikování lze očekávat v důsledku obtížnosti modelování různých variant COVID-19, hustoty zalidnění, logistiky a sociálních aspektů, jako je kultura a životní styl (Chandra et al., 2022). Stežejní záležitostí je komunikování rozsahu možných hodnot v predikcích, přičemž užité statistické koncepty a sama míra nejistoty by měly být snadno interpretovatelné zainteresovanými institucemi a veřejností (Nixon et al., 2022).

## 2.2 Metodika Ministerstva zdravotnictví České republiky

První predikční model od ÚZIS (Ústav zdravotnických informací a statistiky) v rámci metodické dokumentace MZČR popisují Májek et al. (2021a) za účelem krátkodobé simulace počtu nových případů COVID-19. Toho je následně využito k hodnocení dynamiky epidemie, pro stanovení reprodukčního čísla, určování souvislostí se zavedenými opatřeními a jako podkladu pro predikce dalších klíčových charakteristik (Májek et al., 2021a). Jedná se o zjednodušenou aplikaci kompartmentového SIR modelu, v němž jsou všichni jedinci v populaci v každém okamžiku zařazeni do určité stavové skupiny, a jehož parametry byly zvoleny metodou náhodného prohledávání prostoru v přípustném rozsahu hodnot (Májek et al., 2021a).

Koncepčně příbuzný predikční model také od ÚZIS představují Májek et al. (2021b), jehož hlavním účelem je provádět dlouhodobé simulace pro odhad dynamiky epidemie, dopad zavedených opatření, nových variant viru, a také dlouhodobé plánování kapacit zdravotnického systému. Pro implementaci byl adaptován kompartmentový věkově strukturovaný SEIR (SIR + *Exposed*) model, který představili Prem et al. (2020) s veřejně dostupným kódem (URL1) umožňující brát v potaz specifitu prostředí (např. školy, práce, veřejné prostory) spolu s mírou očkovaných jedinců a novými mutacemi viru COVID-19 (Májek et al., 2021b). Vstupem jsou charakteristiky původní studie od Prem et al. (2020), charakteristiky krátkodobého modelu od Májek et al. (2021a), matice kontaktů z mezinárodní, na ČR extrapolované studie od Prem et al. (2017), a vybrané epidemiologické charakteristiky pro ČR (Májek et al., 2021b).

Další uvedený predikční model od ÚZIS popisují Jarkovský et al. (2021), který nachází využití u krátkodobých předpovědí počtu nově a aktuálně hospitalizovaných pacientů, je schopen odlišit hospitalizace obecně od hospitalizací vyžadující intenzivní péči a pohybuje se na úrovni celé ČR a krajů. Za tímto účelem v sobě model kombinuje čtyři samostatné modely: model, který popsal Májek et al. (2021b), klinický model řešící pravděpodobnost hospitalizace od zjištění positivity, kompartmentový model řešící pravděpodobnost úmrtí a propuštění od doby zahájení hospitalizace a model distribuce hospitalizační péče, v němž je na základě dosavadní struktury pacientů odhadnuta očekávaná zátěž jednotlivých nemocnic (Jarkovský et al., 2021).

Poslední uvedený model popisují Příbylová a Hajnová (2021) z Ústavu matematiky a statistiky Přírodovědecké fakulty Masarykovy univerzity v Brně. Představený model ZSEIAR (SEIR + *i*Zolování/Zdraví + *Absent*) rozšiřuje původní SEIR model z hlediska množství dostupných skupin k rozdělení populace o kompartment asymptomatických pozitivních a o skupinu průběžně navyšující velikost náchylné populace (Příbylová

a Hajnová, 2021). Model má díky svému principu limitovanou schopnost dlouhodobé predikce, ale velmi dobrou schopnost monitoringu epidemiologické situace i možnost částečně sledovat rozdíly v jednotlivých krajích, přičemž díky provedené optimalizaci na viru exponované jedince je krátkodobá predikce hospitalizací a smrtí velice dobrá za předpokladu stabilní virulence (Příbylová a Hajnová, 2021).

## 2.3 Akademické práce v České republice

Přestože problematice COVID-19 bylo na českých univerzitách věnováno značné výzkumné úsilí v oblasti metod diagnostiky onemocnění a psychosociálních dopadů na společnost a jedince, nepatří metody predikčního modelování k frekventovaným tématům.

Ve své diplomové práci hodnotí Šálená (2021) užitím SIR modelu vybraná opatření proti šíření COVID-19 s ohledem na jejich maximální potenciál pro ochranu obyvatel ČR, přičemž účinnost těchto opatření posuzovala na základě historických dat z vybraných dříve zasažených zemí. Příbuznou práci poskytuje Matoška (2022), jehož tématem v bakalářské práci byla podrobná analýza některého z dostupných matematických modelů pro predikování vývoje epidemie onemocnění COVID-19 v různých zemích světa. Nebyla nalezena žádná kvalifikační práce aplikující technologie ANN pro potřeby časoprostorového predikčního modelování šíření COVID-19 na území ČR.

V rámci prací Katedry geoinformatiky Přírodovědecké fakulty Univerzity Palackého v Olomouci lze jistou míru příbuznosti nalézt u ostatních prací s tematikou COVID-19, z nichž Kačírková (2022) a Vítková (2022) řeší prostorové analýzy vybraných jevů, Luz (2021) se zabývá designem atlasu s tematikou dopadu pandemie a Porti Suárez (2022) analyzuje uživatelský aspekt čtení datových vizualizací. Metod ML využila Nevτίpilová (2013) k testování kvality prostorových interpolací užitím modulu pro ANN v softwaru GRASS GIS. Dále Tyl (2013) využil architekturu ANN pro shlukování, tzv. *self-organizing map*, pomocí nástroje vytvořeného na Vysoké škole Báňské v Ostravě k analýze výsledků voleb do zastupitelských orgánů a samospráv obcí. Sadílek (2021) nejnověji využil ML metod k identifikaci výtvarného stylu několika sad tematických map pomocí programu Orange.

## 3 METODY A POSTUP ZPRACOVÁNÍ

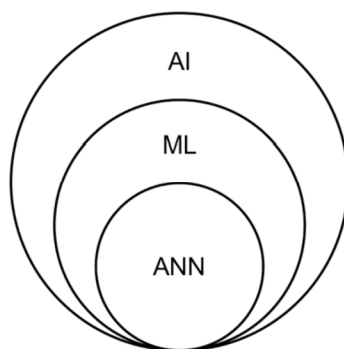
V návaznosti na Široký (2011) lze výzkumný cíl práce považovat za kombinovaný, přičemž funkční složku představuje rozbor vybraných témat tvorby ANN v kontextu predikování, objektovou složku představuje finální výstupní model pro počet aktuálně nakažených COVID-19. Za tímto účelem jsou z obecného hlediska použity specifické vědecké metody matematické, statistické a modelovací.

### 3.1 Použité metody

První fází bylo studium relevantních matematických principů a osvojení si praktických programátorských dovedností potřebných pro implementaci výsledného modelu. Obě složky byly rozšiřovány i během samotné realizace práce.

Vstupu dat do ANN předchází jejich analýza metodami teorie časových řad a statistického testování s následnou transformací do lépe modelovatelné podoby. Zjištěné vlastnosti umožňují optimálněji inicializovat predikční model a usnadňují proces učení sítě. Jmenovitě jde o testování stacionarity, analýzu míry autokorelace a metody datové transformace. Vybranými metodami se zabývá 4. kapitola.

Hlavní téma a jádro práce je vztaženo k velmi širokému a dynamickému oboru umělé inteligence (anglicky *artificial intelligence*, dále AI), jehož zájmem je v dnešní době především tvorba systémů simulujících lidskou inteligenci a intuici u formálně obtížně popsatečných problémů (Goodfellow et al., 2016). ML je jednou z podoblastí AI tvořící metody, které umožňují počítačovým systémům učit se z již existujících zkušeností (tj. dat) za účelem optimalizace jejich chování pro danou úlohu (Nguyen et al., 2019), přičemž je možné nalézt původně zcela neznámá řešení (Joshi, 2023). ANN představují podmnožinu ML metod inspirovanou biologickými neuronovými sítěmi živých organismů (Nguyen et al., 2019). Vztah mezi AI, ML a ANN je schematizován na Obr. 1. Vybranými metodami ML a ANN se zabývá podkapitola 5.1 obsahující rozbor relevantní teorie pro výslednou praktickou implementaci. Způsob této implementace užitím autorem práce sestaveným nástrojem je představen v podkapitole 5.2.



Obr. 1 Vztah mezi AI, ML a ANN.

Z povahy procesu tvorby ANN plynula potřeba dlouhodobého testování různě strukturovaných modelů ve snaze komparativní metodou nalézt co nejoptimálnější řešení, které je představeno v 6. kapitole. Její součástí je také samotný postup evaluace modelu a finální vizualizace výsledků kartografickými metodami v souladu s Voženílek et al. (2011).

## 3.2 Použitý software a hardware

K realizaci vědeckých metod práce je použit svobodný a otevřený software, což dělá celý proces dále snadněji replikovatelný a modifikovatelný pro další podobné úlohy. Díky tomu roste potenciál pro navazující výzkum a využitelnost představených postupů.

Přehled užitých softwarových nástrojů poskytuje *Tabulka 1*.

*Tabulka 1* Přehled užitých softwarových prostředků a jejich rozšíření.

	<b>Verze</b>	<b>Účel v práci</b>
<b>Python</b>	3.7.9	programovací jazyk
BLiTZ	0.2.8	bayesovské vrstvy pro PyTorch
Matplotlib	3.5.3	datové vizualizace
NumPy	1.19.1	efektivní práce s vícerozměrnými poli
Optuna	2.10.0	automatická optimalizace hyperparametrů
Pandas	1.1.1	import a restrukturalizace vstupních dat, export
PyTorch	1.10.2+cu102	tvorba ANN a práce s tenzory
Statsmodels	0.13.2	statistické testování a analýza dat
Sklearn	0.23.2	transformace vstupních dat
tqdm	4.64.0	ukazatel iteračních procesů
<b>QGIS</b>	3.16.3	tvorba map, finální výpočty a analýza
Visualist	0.6	analýza prostorového shlukování
<b>CUDA</b>	12.0	zrychlení procesu učení ANN

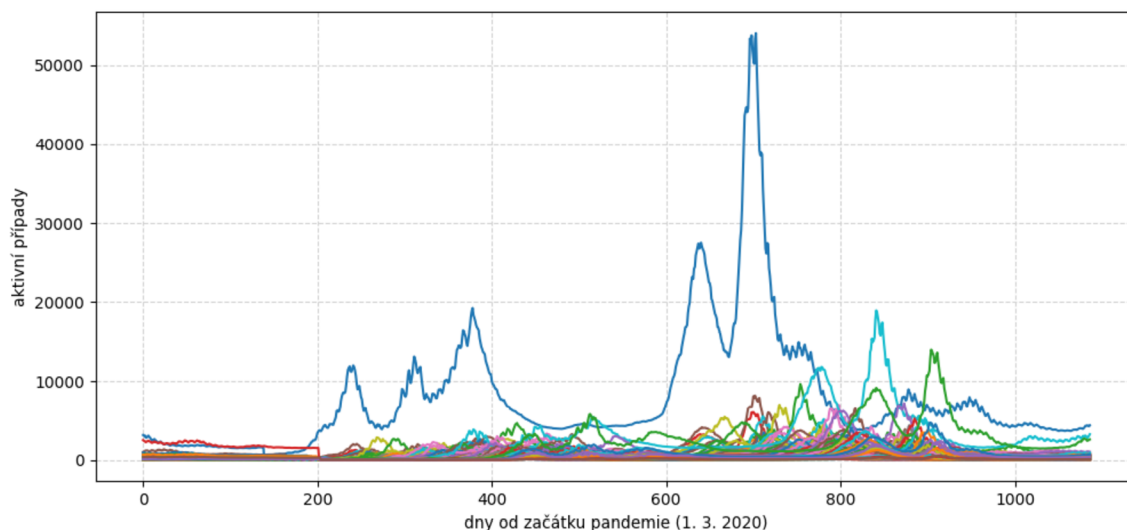
Knihovna PyTorch umožňuje snadnou interoperabilitu s technologií CUDA (*Compute Unified Device Architecture*), jejímž užitím je možné zrychlit výpočetně náročné operace při učení ANN přesunem na GPU (*Graphics Processing Unit*) (Nguyen et al., 2019). K tomuto účelu je použito GPU Nvidia GeForce GTX 1070 Ti.

## 3.3 Použitá data

Epidemiologická data pochází z otevřených datových sad *Přehled epidemiologické situace dle hlášení krajských hygienických stanic podle ORP (obec s rozšířenou působností) a Epidemiologická charakteristika městských částí hlavního města Prahy* od ÚZIS z COVID-19 portálu MZČR (URL2). Epidemiologickými charakteristikami je tak pokryto všech 205 ORP a Praha, tedy celé území ČR.

Za modelovaný jev byl zvolen aktuální počet nakažených s denní frekvencí, který lze následně přepočítat na prevalenci, tj. proporce osob mající daný klinický stav v určitém čase, s významem pro plánování a hodnocení potřeby zdravotní péče (Janout a Janoutová, 2021).

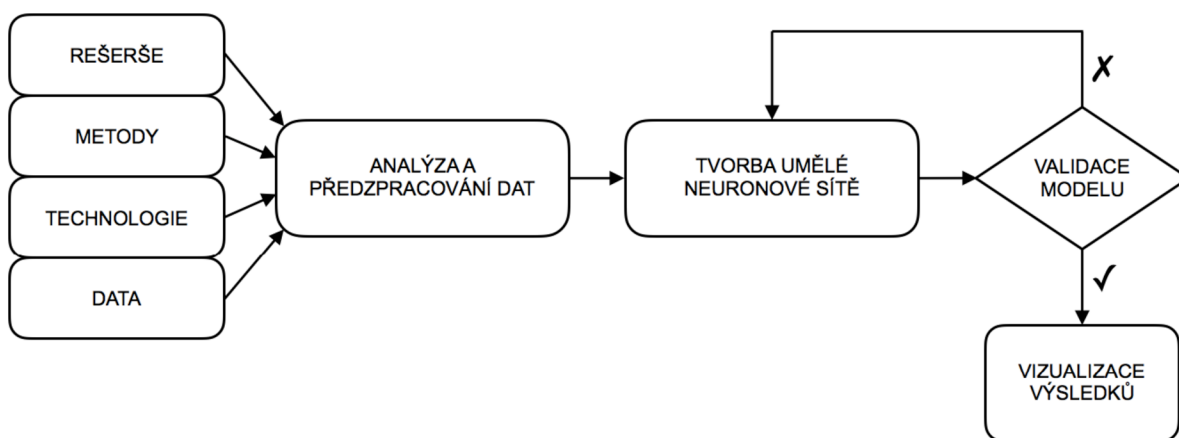
Po restrukturalizaci a sloučení datových sad (užitím knihovny Pandas viz *Příloha 2*) mají vstupní data modelu podobu 206 paralelních časových řad o délce 1087 dnů (období 1. 3. 2020–21. 2. 2023). Vizualizována jsou na *Obr. 2*.



Obr. 2 Vstupní data modelu.

Pro mapové vizualizace je využita prostorová datová sada ArcČR 4.1 (URL3), konkrétně vrstva hranic ORP a její atributové informace z SLDB (Sčítání lidu, domů a bytů) 2021. Propojení epidemiologických a prostorových dat je realizováno přes kódovou identifikaci daných území.

### 3.4 Postup práce



Obr. 3 Schéma postupu práce.



## 4 ANALÝZA A PŘEDZPRACOVÁNÍ VSTUPNÍCH DAT

Mnoho případů využití ANN vyžaduje sofistikované předzpracování dat, jelikož reálná vstupních data mohou mít ve své surové podobě pro mnoho sítí obtížně zpracovatelné vlastnosti (Goodfellow et al., 2016). Snahou je využít sekvenční povahu vstupních dat k získání hodnotných informací k objektivnější inicializaci ANN a k následné transformaci dat do lépe zpracovatelné podoby.

### 4.1 Stacionarita

Stacionarita časové řady je stanovena ve vztahu k proměnlivosti statistických vlastností v čase a její pozitivní výskyt bývá obvykle charakterizován jako konstantní rozdělení pravděpodobnosti po celé délce sekvence dat, což lze v principu vnímat tak, že libovolně zvolené časové výseče sledované veličiny budou mít podobné popisné charakteristiky (Montgomery et al., 2015).

Jeden z často používaných statistických testů k robustnější validaci stacionarity je ADF (*Augmented Dickey-Fuller*) test. Ten představili Said a Dickey (1984) v návaznosti na Dickey a Fuller (1979) k vyhodnocení přítomnosti tzv. jednotkového kořene (anglicky *unit root*) pro danou časovou řadu (Herranz, 2017). Stručné znění nulté hypotézy  $H_0$  a alternativní hypotézy  $H_A$  poskytuje dokumentace knihovny Statsmodels (URL4):

$H_0$ : Časová řada obsahuje jednotkový kořen (není stacionární),

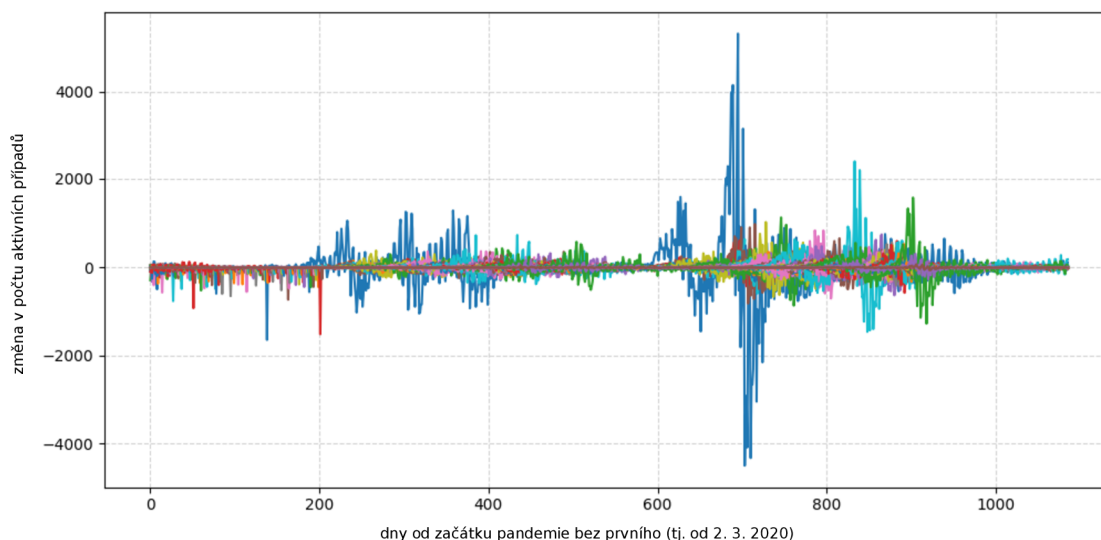
$H_A$ : Časová řada neobsahuje jednotkový kořen (je stacionární).

Podrobnosti týkající se jednotkového kořene a exaktního vyhodnocování ADF testu přesahují možnosti této práce a poskytuje je například Herranz (2017), s důkazy dále Said a Dickey (1984).

Je-li potvrzena  $H_0$ , lze data přesto udělat stacionární tzv. diferencováním (Said a Dickey, 1984), což odpovídá výpočtu rovnice, upravené z Montgomery et al. (2015):

$$\forall t \in T: \Delta x_t = x_t - x_{t-1} \mid \Delta x_1 = \emptyset \quad [4.1]$$

kde  $T$  je  $t$ -prvková časová řada jevu  $x$  a  $\Delta$  označuje diferencovanou hodnotu. Jelikož pro první prvek sekvence neexistuje žádný předcházející, je tento prvek vypuštěn, což implikuje zkrácení délky časové řady z  $T$  na  $T - 1$ , přičemž pokud data nejsou stacionární ani po diferencování, lze celý proces opakovat a obdržet tak diferenci 2. řádu, případně vyššího (Hyndman a Athanasopoulos, 2021). Efekt difference 1. řádu na vstupní data z Obr. 2 poskytuje Obr. 4.



Obr. 4 Diferencovaná vstupní data.

Užití LSTM (a tedy i GRU) je preferováno pro složité časové řady, jelikož jsou schopny naučit se nelineární a nestacionární data pro potřeby predikcí (Patarwal et al., 2019). Na druhou stranu kupříkladu Virili a Freisleben (2000) zastávají spíše kritické stanovisko, protože nestacionární povaha predikovaných sekvencí vede u mnoha případů k výraznému snížení přesnosti, ačkoliv konstatují, že by tomu tak u ANN teoreticky být nemělo. Debatu ohledně schopnosti ANN zpracovávat dostatečně přesně nestacionární data lze tak zatím považovat za otevřenou.

## 4.2 Autokorelace

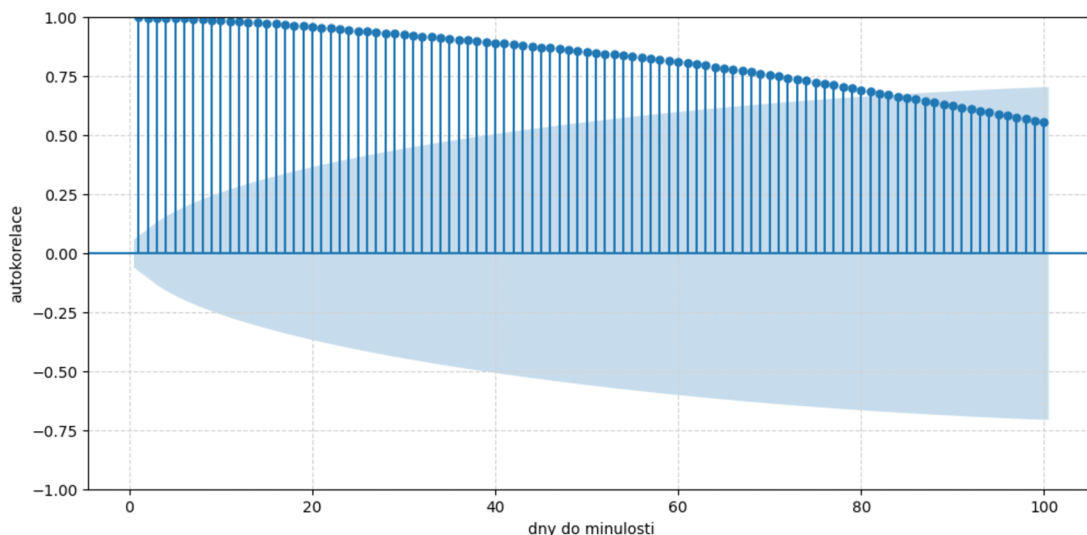
Podobně jako korelace popisuje lineární závislost mezi dvěma proměnnými, pomocí níž je možné sestavovat regresní modely, měří autokorelace lineární závislost mezi jednotlivými prvky časové řady pro potřeby tvorby autoregresních modelů, v nichž jsou budoucí hodnoty predikovány na základě hodnot předcházejících (Hyndman a Athanasopoulos, 2021).

Jedna z metod určení míry autokorelace je ACF (*Autocorrelation Function*), jejímž výstupem je množina hodnot získaných výpočtem rovnice autokorelačního koeficientu rozšířeného z Hyndman a Athanasopoulos (2021) podle Montgomery et al. (2015):

$$r_k = \frac{cov(x_t, x_{t+k})}{var(x_t)} = \frac{\sum_{t=k+1}^T (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2} \quad [4.2]$$

kde  $r_k$  označuje hodnotu autokorelačního koeficientu pro prvek časové řady  $T$  na pozici  $t$  vůči  $k$  předcházejícím prvkům sekvence,  $cov$  označuje kovarianci,  $var$  rozptyl,  $x$  hodnoty sledovaného jevu a  $\bar{x}$  jeho průměrnou hodnotu.

Vizualizace hodnot ACF spolu s intervaly spolehlivosti o dané hladině spolehlivosti  $\alpha$  následně umožňuje určit statisticky významné hodnoty autokorelace (Homayouni et al., 2020), což je snadno implementovatelné díky knihovně Statsmodels. Na Obr. 5 je k dispozici graf ACF pro průměr časových řad z Obr. 2.



Obr. 5 Hodnoty ACF průměru vstupních dat pro  $\alpha = 0,05$ .  
(dny s hodnotami mimo modrou plochu jsou statisticky relevantní)

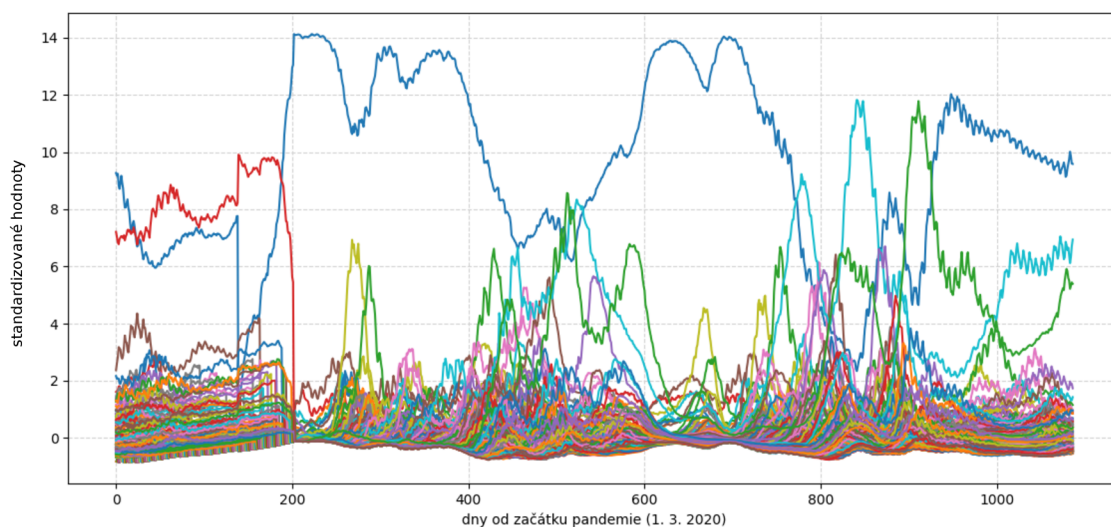
Alternativním určením míry autokorelace je PACF (*Partial ACF*), která sice dosahuje velmi podobných výsledků, ale na rozdíl od ACF odstraňuje interferenci mezilehlých hodnot v sekvenci (Flores et al., 2012).

Podle Zhang et al. (1998) není autokorelace z důvodu své lineární povahy přímo převoditelná do oblasti ANN, jelikož ty modelují v zásadě nelineární procesy. Na druhou stranu kupříkladu Homayouni et al. (2020) využili ACF při rozhodování o zvolené délce tzv. posuvného okna k dosažení velmi dobrých výsledků. Tato metoda je použita i v praktické realizaci cílů této práce (viz podkapitola 5.2).

### 4.3 Datové transformace

Pro většinu praktických aplikací je nutné původní data transformovat do nové podoby tak, aby bylo rozpoznávání vzoru v datech pro ML algoritmus snadnější (Bishop, 2006). Často používanými přístupy jsou standardizace, tj. transformace dat do podoby s nulovým průměrem a jednotkovou směrodatnou odchylkou, a normalizace, tj. transformace dat do stanoveného intervalu pomocí minimální a maximální hodnoty (Ali a Faraj, 2014). Povaha modelovaných COVID-19 dat nicméně normalizaci vylučuje, jelikož lze pozorovat značný rozdíl v absolutním počtu aktuálně nakažených mezi jednotlivými oblastmi a také abnormální hodnoty v Praze.

Standardizace průměrem a směrodatnou odchylkou předpokládá normální rozdělení transformovaných dat (Aggarwal, 2018). Efekt této standardizace na vstupní data z Obr. 2 poskytuje Obr. 6, doplněný o výpočet podle URL5:

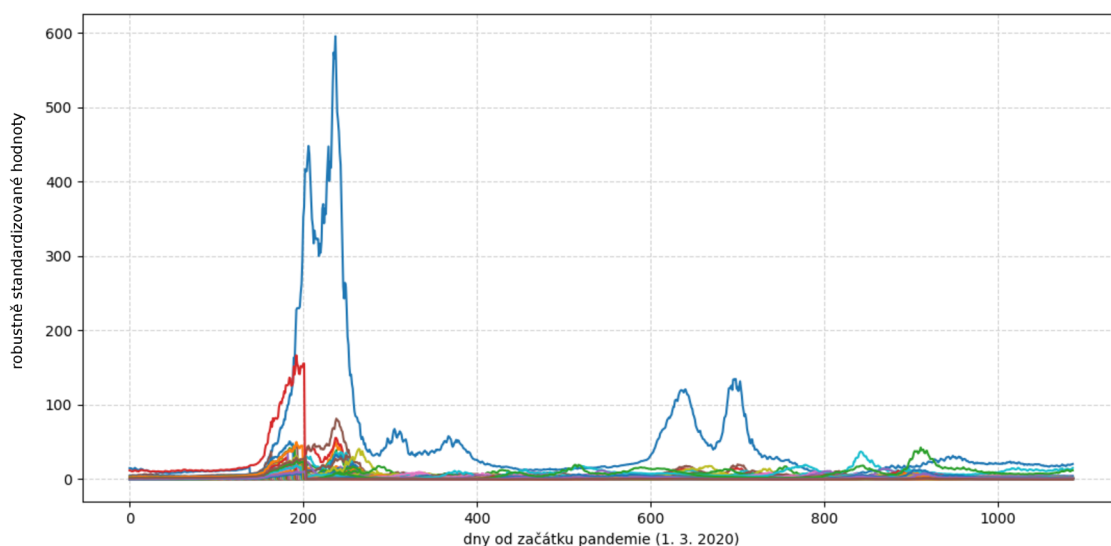


$$x_s = \frac{x - \bar{x}}{\sigma} \quad [4.3]$$

Obr. 6 Standardizace vstupních dat.

kde  $x_s$  je standardizovaná hodnota jevu  $x$ , jeho průměr je  $\bar{x}$  a  $\sigma$  je směrodatná odchylka.

Pokud podmínka normality není zajištěna a data obsahují výrazně odlehle hodnoty, je vhodnější použít robustnější metodu standardizace mediánem a mezikvartilovým rozpětím (URL6). Efekt této robustní standardizace na vstupní data z Obr. 2 poskytuje Obr. 7, doplněný výpočtem podle URL7:

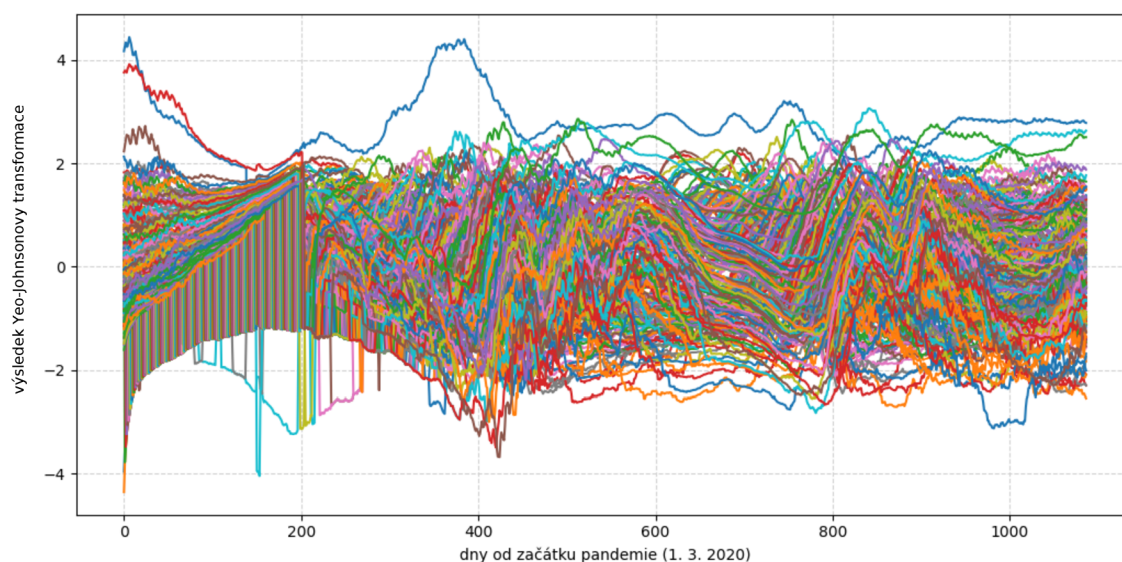


$$x_s = \frac{x - \tilde{x}}{Q_{0,75} - Q_{0,25}} \quad [4.4]$$

Obr. 7 Robustní standardizace vstupních dat.

kde  $x_s$  je robustně standardizovaná hodnota jevu  $x$ , jeho medián je  $\tilde{x}$ ,  $Q_{0,75}$  je 3. kvartil a  $Q_{0,25}$  je 1. kvartil, tedy  $Q_{0,75} - Q_{0,25}$  je mezikvartilové rozpětí.

Velmi užitečné jsou také tzv. mocninné transformace, tj. parametrické monotónní nelineární transformace, mající za cíl udělat vstupní data bližší normálnímu rozdělení (URL8). Pro potřeby této práce se za vhodnou prokázala Yeo-Johnsonova transformace, kterou představili Yeo a Johnson (2000), umožňující transformovat i nulové a záporné hodnoty. Efekt této standardizace na vstupní data z Obr. 2 poskytuje Obr. 8, doplněný výpočtem podle URL9:



$$x^{(\lambda)} = \begin{cases} \frac{(x+1)^\lambda - 1}{\lambda} & | \lambda \neq 0 \wedge x \geq 0 \\ \ln(x+1) & | \lambda = 0 \wedge x \geq 0 \\ -\frac{(-x+1)^{2-\lambda} - 1}{2-\lambda} & | \lambda \neq 2 \wedge x < 0 \\ -\ln(-x+1) & | \lambda = 2 \wedge x < 0 \end{cases} \quad [4.5]$$

Obr. 8 Yeo-Johnsonova transformace dat.

kde  $x^{(\lambda)}$  označuje transformovaná data  $x$  parametrem  $\lambda$ , který je algoritmicky určen.

Z porovnání grafů plyne jistá výhoda Yeo-Johnsonovy transformace oproti oběma standardizacím, jelikož na rozdíl od nich úspěšně ztlumila výrazné rozdíly v počtu aktuálně nakažených napříč jednotlivými ORP a také abnormální hodnoty v Praze, které by mohly mít negativní vliv na učení sítě.

Všechny výše zmíněné i mnohé další transformace jsou součástí knihovny Sklearn v rámci níž je možné jednoduše provést i transformace inverzní. Jejich srovnání je dostupné na URL10.

## 5 TVORBA UMĚLÉ NEURONOVÉ SÍTĚ

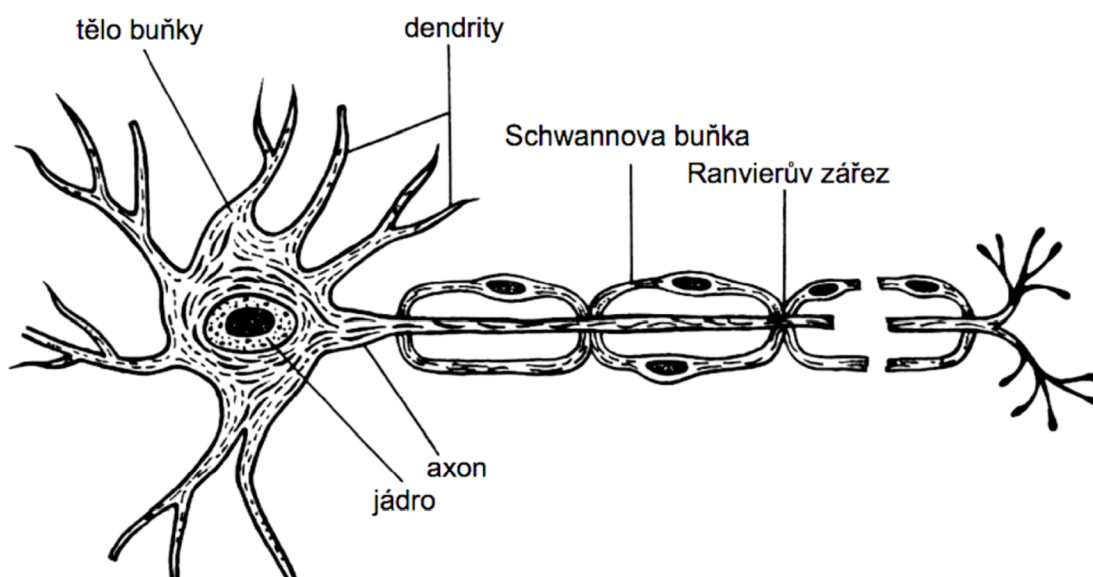
ANN lze charakterizovat jako výpočetní modely s naučitelnými parametry, jejichž dílčí výpočetní vrstvy jsou mezi sebou propojeny jakožto orientované grafy, avšak detailní mechanismy především tzv. hlubokých (anglicky *deep*) ANN (viz podkapitola 5.1.3) nejsou zatím jednoznačně matematicky uchopeny (Kaul a Lall, 2020). To spolu s jejich mnoha úspěšnými aplikacemi může vést výzkumníky ke zvýšenému úsilí o jejich výzkum, chtějí-li se držet na popředí současného vývoje AI (Kruse et al., 2022).

Jeden z hlavních aspektů tvorby ANN je stanovení její architektury (tj. množství, typ propojení a celková konfigurace dílčích neuronových vrstev), které bývají obvykle specializovány na specifické úkoly (Goodfellow et al., 2016). V konečném důsledku však tyto architektury mají obvykle širší aplikovatelnost, což spolu s obtížností stanovení účinnosti konkrétních řešení vede k potřebě rigorózního testování kandidátních modelů sestavovaných s ohledem na povahu vstupních dat a účel finalizovaného modelu (Nguyen et al., 2019).

### 5.1 Relevantní teoretické principy

Základní složky a koncepty ANN jsou inspirovány anatomii a fyziologií biologických neuronových sítí, které v nervovém systému živočichů plní roli hlavního zpracovatele informací (Kruse et al., 2022). Přestože jakožto matematické simulace jsou fundamentálně odlišné od fyzikálně-chemického fungování biologických neuronů, zůstává vývoj ANN a nové objevy v neurologii vzájemně propojeny (Aggarwal, 2018).

Konceptuální podobnost mezi umělým a biologickým neuronem je patrná ze srovnání Obr. 9 a Obr. 11. Perspektiva neurofyziologie přesahuje možnosti této práce a pozornost jí věnují například Kittnar et al. (2020), detailněji pak Mills (2017).



Obr. 9 Základní stavba biologického neuronu (Seidl, 2015).

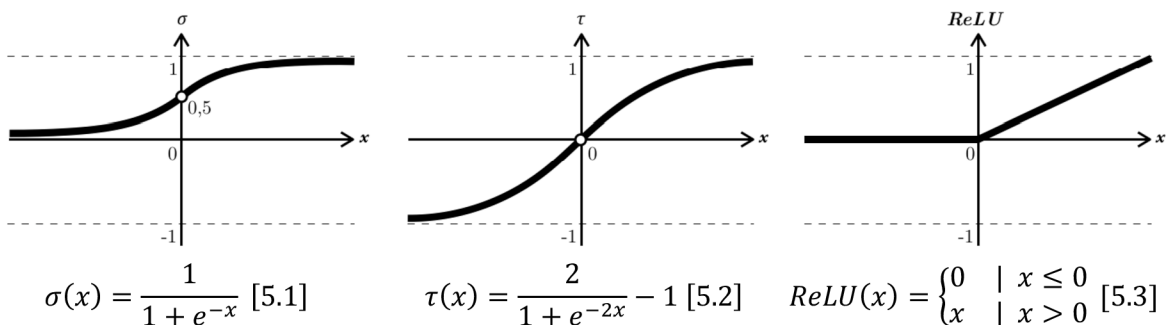
Kapitola stručně představuje pouze relevantní metody ML a ANN pro finální praktickou implementaci (viz podkapitola 5.2). Rozbor dalších metod, byť stejně významných a potenciálně využitelných, přesahuje možnosti této práce.



### 5.1.1 Umělý neuron

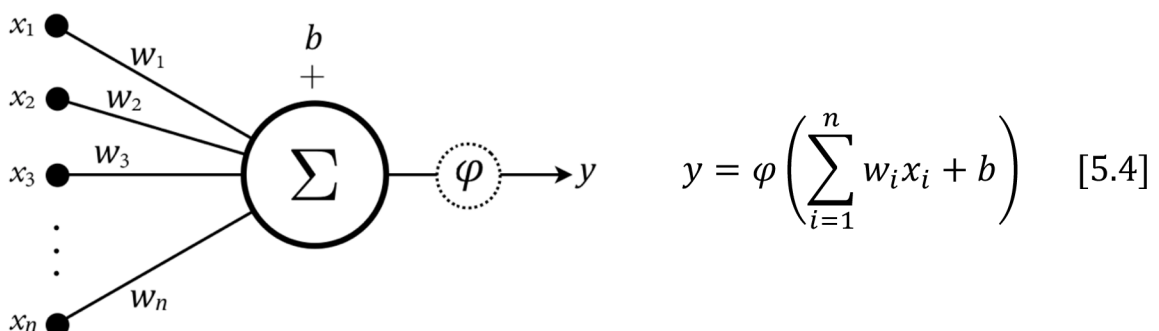
McCulloch a Pitts (1943) představili historicky první matematický model neuronu, který vycházel přímo z fyziologického fenoménu akčního potenciálu a sloužil k řešení logických operací, jelikož na základě binárních vstupů neuron na výstupu buď zahorel (překonal prahovou hodnotu), nebo nezahorel (nepřekonal ji) (Kruse et al., 2022). Na to navázal Rosenblatt (1958) představením tzv. perceptronu schopného přijímat vážené numerické vstupy, na jejichž základě prováděl binární klasifikaci (Aggarwal, 2021).

Finální výstup neuronu úzce souvisí s problematikou tzv. aktivačních funkcí (anglicky *activation functions*), jimiž je transformován do určitého intervalu (Haykin, 2009). V tomto smyslu oba výše zmíněné modely využívají tzv. prahovou funkci (anglicky *threshold function*), nicméně u moderních ANN ji nahrazují např. sigmoidální funkce, hyperbolický tangens, *rectified linear unit*, *softmax* atd. (Kruse et al., 2022). Volba aktivačních funkcí je zásadní složkou tvorby ANN, jelikož užití každé z nich je vázáno k určitému typu modelu (Aggarwal, 2018). Další vlastnosti aktivačních funkcí jsou zmíněny v podkapitolách 5.1.3 a 5.1.8, kdy budou lépe kontextualizovány. Na Obr. 10 jsou znázorněny často používané aktivační funkce včetně jejich matematického předpisu.



Obr. 10 Často používané aktivační funkce, upraveno ze Zaki a Meira (2020).  
(zleva: sigmoidální funkce, hyperbolický tangens, *rectified linear unit*)

Umělý neuron lze v souladu s výše stanoveným schematizovat na Obr. 11 uvedeným způsobem:



Obr. 11 Schéma a výpočet umělého neuronu, upraveno z Haykin (2009).

kde  $y$  je výstup neuronu,  $\varphi$  je aktivační funkce,  $w_1, \dots, w_n$  jsou váhy vstupů,  $x_1, \dots, x_n$  jsou vstupy a  $b$  je tzv. *bias*, vzniklý přesunem prahové hodnoty na levou stranu v původní nerovnici perceptronu (Kruse et al., 2022).

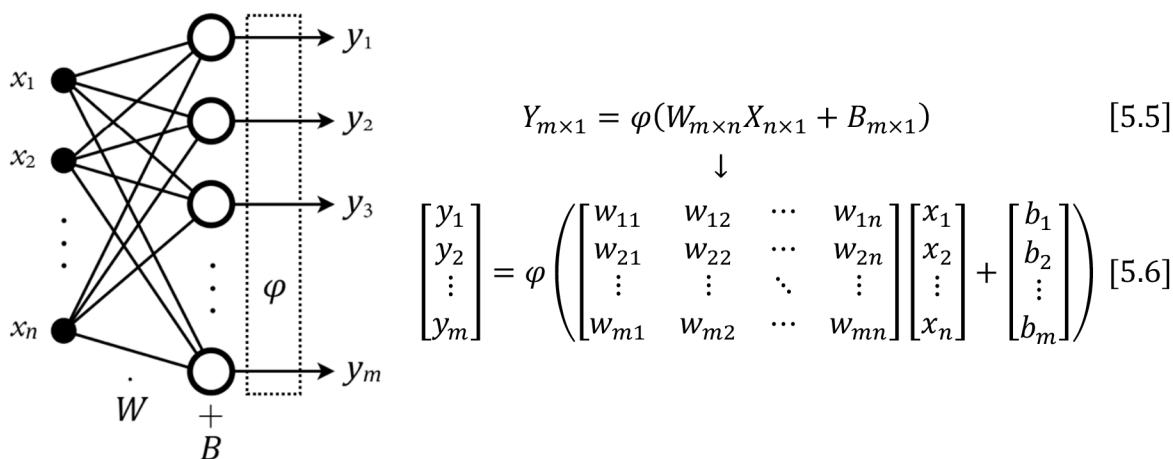
Aktivační funkci  $\varphi$  je sice možné vynechat, výsledný výstup však bude identický jako po užití tzv. lineární aktivační funkce (Aggarwal, 2018). Z tohoto důvodu je aktivační funkce uvedena jako integrální součást popisu neuronu.

### 5.1.2 Vrstva umělých neuronů

Seskupením více umělých neuronů vzniká vrstva umělých neuronů, umožňující obecně projekci  $n$  vstupů na  $m$  výstupů, přičemž vstupy a výstupy jsou ve výchozí formě vzájemně propojeny každý s každým (Haykin, 2009).

Stručný zápis výpočtu umožňují nástroje lineární algebry, jejichž principy jsou zcela zásadní pro pochopení a využití mnoha ML algoritmů (Goodfellow et al., 2016). Formy matic a vektorů je využíváno i při počítačové implementaci ANN, jak ilustruje dokumentace knihovny PyTorch kupříkladu při popisu základní vrstvy (URL11).

Vrstvu umělých neuronů lze v souladu s výše stanoveným schematizovat způsobem uvedeným na Obr. 12:



Obr. 12 Schéma a výpočet vrstvy umělých neuronů, upraveno z Haykin (2009) a rozšířeno.

kde  $Y$  je vektor výstupů,  $\varphi$  je aktivační funkce,  $W$  je náhodně (nebo sofistikovaněji) inicializovaná matice vah,  $X$  je vektor vstupů a  $B$  je vektor *biases*. Matice a vektory budou nadále nahrazeny souborným označením tenzor popisující obecně libovolné  $d$ -rozměrné pole (Goodfellow et al., 2016), což je stručné a účelné zároveň, jelikož s tenzory pracuje i knihovna PyTorch (URL12).

### 5.1.3 Umělá neuronová síť

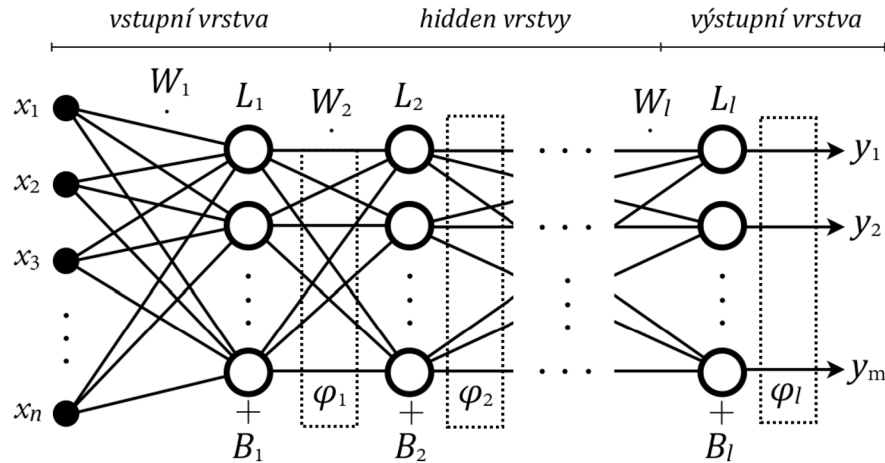
Napojením dvou a více více neuronových vrstev za sebe vzniká vícevrstevná neuronová síť, jejíž základní varianta předpokládá úplné, tzv. *dense* propojení neuronů sousedících vrstev (Aggarwal, 2018). Vrstvy bez přímého spojení s tenzory  $X$  nebo  $Y$  jsou označovány jako skryté (anglicky *hidden*), a svými přidávanými synapsemi umožňují ANN efektivněji extrahovat informace ze vstupu (Haykin, 2009). ANN obsahující více než jednu *hidden* vrstvu jsou označovány jako tzv. *deep* ANN, které lze v jejich principu považovat za nejúčinnější ML algoritmus pro široké spektrum úloh (Kaul a Lall, 2020).

Libovolně hluboká ANN používající pouze lineární (ekvivalentně také žádné) aktivační funkce je stejně expresivní jako pouze jedna její vrstva (Kruse et al., 2022). To lze vyložit tak, že nelineární aktivační funkce na *hidden* vrstvách provádí nelineární transformace prostoru vstupních dat (anglicky *input space*) do tzv. *feature space* za účelem jeho



zjednodušené a snadněji uchopitelné reprezentace, přičemž tato schopnost roste s množstvím nelineárně aktivovaných *hidden* vrstev v síti (Haykin, 2009). Animaci příkladu této transformace lze nalézt na URL13.

Umělou neuronovou síť lze v souladu s výše stanoveným schematizovat způsobem uvedeným na Obr. 13:



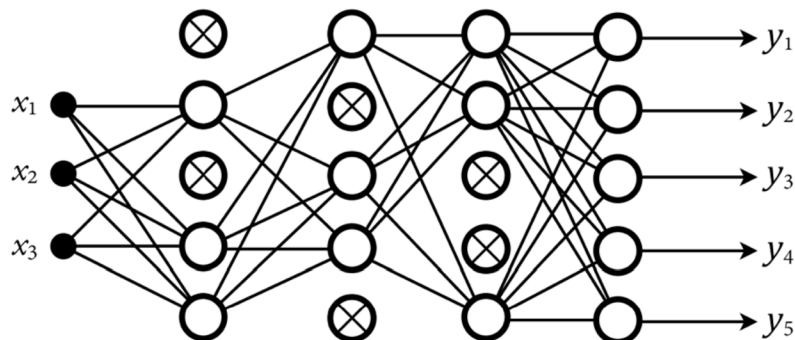
Obr. 13 Schéma  $l$ -vrstvé ANN, upraveno z Aggarwal (2021) a rozšířeno.

kde  $x_1, \dots, x_n$  jsou vstupy,  $y_1, \dots, y_m$  výstupy,  $L_1, \dots, L_l$  vrstvy ANN,  $W_1, \dots, W_l$  tenzory vah,  $B_1, \dots, B_l$  tenzory biases a  $\varphi_1, \dots, \varphi_l$  aktivační funkce. Všem *hidden* vrstvám a výstupní vrstvě bylo pro přehlednost uděleno shodně  $m$  neuronů, nicméně ve skutečnosti je možné mít každou vrstvu jinak širokou.

Navazující pohled na ANN jako na sérii vnořených funkcí je představen v podkapitole 5.1.8, jelikož přímo souvisí s problematikou učení.

### 5.1.4 Dropout

Metoda *dropout* spočívá v náhodné deaktivaci (resp. stanovení nulových vah a *biases*) stanoveného procenta neuronů ve vstupní nebo *hidden* vrstvě ve fázi učení pro každou epochu, čímž efektivně snižuje riziko přeučení (anglicky *overfitting*, viz podkapitola 5.1.8) (Kruse et al., 2022). Metodu nelze využít u výstupní vrstvy, jelikož by nebylo možné spolehlivě počítat ztrátovou funkci (viz 5.1.7) (Aggarwal, 2018). *Dropout* lze také vnímat jako jednu z tzv. *ensemble* technik skládající dílčí podsítě vzniklé randomizovanými deaktivacemi do jedné finální, přičemž je výpočetně úsporný a využitelný takřka všemi typy ANN (Goodfellow et al., 2016). Princip metody je znázorněn na Obr. 14.

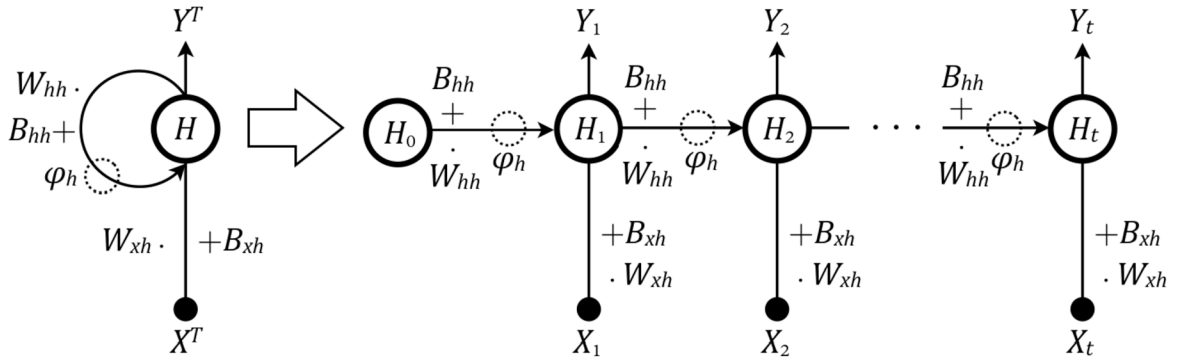


Obr. 14 Princip *dropout* (zde 40 % pro všechny dostupné vrstvy), upraveno z Kruse et al. (2022).

### 5.1.5 Rekurentní neuronové vrstvy

ANN se stává RNN, obsahuje-li alespoň jednu zpětnovazebnou smyčku (anglicky *feedback loop*) (Haykin, 2009). Té je využito v rekurentních neuronových vrstvách k modelování sekvenčních závislostí časových řad, čehož nelze efektivně dosáhnout základní architekturou ANN (Aggarwal, 2021). Rekurentní vrstvy zpracovávají postupně jednotlivé časové kroky vstupní sekvence, a kromě vstupu pro daný okamžik rekurzivně přijímají a produkují i tzv. *hidden state* tenzor, který v sobě implicitně uchovává informace z předcházejících časových kroků (Bianchi et al., 2017).

Základní rekurentní vrstvu lze v souladu s výše stanoveným schematizovat na Obr. 15 uvedeným způsobem:



$\forall t \in T:$

$$H_t = \varphi_h(W_{xh}X_t + B_{xh} + W_{hh}H_{t-1} + B_{hh}) \quad [5.7]$$

$$Y_t = H_t \quad [5.8]$$

Obr. 15 Schéma základní rekurentní vrstvy, upraveno z Aggarwal (2018) podle URL14.

kde  $T$  je časová řada,  $t$  jsou jednotlivé časové kroky  $T$ ,  $H$  je *hidden state* tenzor,  $X^T$  a  $Y^T$  jsou tenzory vstupů a výstupů se sekvencemi o délce  $T$ ,  $W_{xh}$  a  $B_{xh}$  jsou tenzory vah a *biases* mezi  $X^T$  a  $H$ ,  $W_{hh}$ ,  $B_{hh}$  a  $\varphi_h$  jsou tenzory vah, *biases* a aktivační funkce v rámci rekurzivního výpočtu  $H$ . Za  $\varphi_h$  bývá typicky volen hyperbolický tangens nebo *rectified linear unit* a  $H_0$  bývá obvykle inicializován jako tenzor nul (URL14).

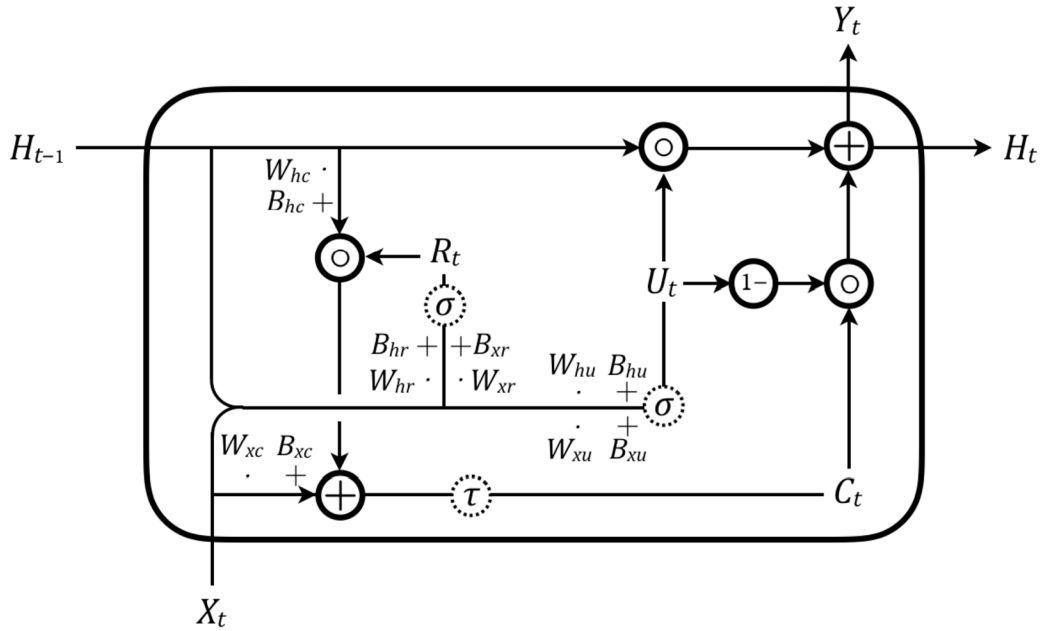
Tenzory vah a *biases* zůstávají stejné po celou dobu průchodu sekvencí  $T$  (Bianchi et al., 2017). Z toho plyne jistá míra nestability, jelikož při rekurentním násobení tenzory vah může dojít k tzv. *vanishing* (tenzory vah obsahují příliš malé hodnoty) nebo *exploding* gradientu (tenzory vah obsahují příliš velké hodnoty), což v obou případech vede k fenoménu analogickému s krátkodobou pamětí, jelikož negativní dopad na učení roste s délkou sekvence (Aggarwal, 2021). Role gradientů je blíže vysvětlena v podkapitole 5.1.8.

Řešení tohoto problému představili Hochreiter a Schmidhuber (1997) ve formě modifikované rekurentní vrstvy zvané LSTM, pomocí níž lze modelovat libovolně dlouhé sekvence (Kruse et al., 2022). Modernější alternativa pro LSTM je vrstva GRU vytvořená Cho et al. (2014), která i přes jednodušší výpočetní procesy poskytuje podobný výkon jako LSTM, a dokonce ji překonává na menších datasetech (Kruse et al., 2022).

Základní myšlenkou GRU je systematické aktualizování hodnot *hidden state* tenzoru, k čemuž jsou využity tzv. *update* a *reset gates*, pomocí nichž se reguluje informace z předcházejících časových kroků (Aggarwal, 2018). *Update gate* řeší integraci aktuálního

vstupu sekvence do *hidden state* a *reset gate* řeší resetování, resp. zapomínání určitých hodnot v *hidden state* (Bianchi et al., 2017).

GRU lze v souladu s výše stanovenými poznatky schematizovat na Obr. 15 uvedeným způsobem:



$\forall t \in T$ :

$$R_t = \sigma(W_{xr}X_t + B_{xr} + W_{hr}H_{t-1} + B_{hr}) \quad [5.9]$$

$$U_t = \sigma(W_{xu}X_t + B_{xu} + W_{hu}H_{t-1} + B_{hu}) \quad [5.10]$$

$$C_t = \tau(W_{xc}X_t + B_{xc} + R_t \circ (W_{hc}H_{t-1} + B_{hc})) \quad [5.11]$$

$$H_t = (1 - U_t) \circ C_t + U_t \circ H_{t-1} \quad [5.12]$$

$$Y_t = H_t \quad [5.13]$$

Obr. 16 Schéma GRU, upraveno z Bianchi et al. (2017) podle URL15.

kde  $R_t$  je výstupní tenzor *reset gate*,  $U_t$  je výstupní tenzor *update gate*,  $C_t$  je tzv. *candidate state* tenzor,  $\sigma$  je sigmoidální funkce,  $\tau$  je hyperbolický tangens,  $W_{xr}$ ,  $W_{xu}$ ,  $W_{xc}$ ,  $W_{hr}$ ,  $W_{hu}$ ,  $W_{hc}$ ,  $B_{xr}$ ,  $B_{xu}$ ,  $B_{xc}$ ,  $B_{hr}$ ,  $B_{hu}$ ,  $B_{hc}$ , jsou tenzory vah a *biases* v dílčích fázích průchodu GRU vstvou a symbol  $\circ$  označuje Hadamardův součin, tj. operaci násobení odpovídajících si prvků dvou tenzorů se stejnými rozměry (Goodfellow et al., 2016).

### 5.1.6 Bayesovská parametrizace vrstev

Výše uvedené kapitoly popisovaly deterministické ANN, jejichž parametry  $\theta$  (tj. váhy  $W$  a *biases*  $B$ ) měly podobu konstant. Oproti tomu stochastické BNN (*Bayesian Neural Network*) reprezentují parametry jako náhodné veličiny, jejichž variabilita reflektuje variabilitu v modelovaných datech, přičemž generování několika náhodných výběrů z jejich rozdělení pravděpodobnosti umožňuje vytvářet více podobných výstupů a pracovat s nimi obdobně jako s jinými *ensemble* technikami (Blundell et al., 2015), díky čemuž také snižují riziko přeučení (Jospin et al., 2022).

Optikou bayesovské statistiky můžeme pozorovat vztah mezi parametry a trénovacími daty, který je proporční k předpokladům o samotných parametrech a volbě modelu (Goan a Fookes, 2020). Adaptací bayesovy věty k popisu této skutečnosti se zabývají Goan a Fookes (2020) a jejich závěry je možné sloučit a upravit jako:

$$P(\theta|D_t) = \frac{P(\theta)P(D_t|\theta)}{P(D_t)} = \frac{P(\theta) \prod_{i=1}^n P(Y_i|X_i, \theta)}{\int P(\theta)P(D_t|\theta) d\theta} \quad [5.14]$$

kde  $\theta$  je množina parametrů BNN,  $D_t$  je trénovací množina dat (viz podkapitola 5.1.8),  $P(\theta|D_t)$  je pravděpodobnost *aposteriori*,  $P(\theta)$  je pravděpodobnost *apriori*,  $P(D_t|\theta)$  je tzv. *likelihood*,  $P(D_t)$  je tzv. *evidence* a  $P(Y_i|X_i, \theta)$  po rozepsání  $P(D_t|\theta)$  reprezentuje vliv stanovené architektury BNN prostřednictvím výstupů  $Y_1, \dots, Y_n$  na základě vstupů  $X_1, \dots, X_n$  z  $D_t$  a parametrů sítě  $\theta$ .

Znalost  $P(\theta|D_t)$  umožňuje stanovovat predikce jakožto očekávané hodnoty na základě předchozí zkušenosti s  $D_t$  (Jospin et al., 2022),  $P(\theta)$  reprezentuje předem stanovené očekávání o rozložení parametrů sítě  $\theta$  (Goan a Fookes, 2020), které může být v knihovně BLiTZ určeno jako vážená směs dvou normálních rozdělení nebo jako jiné rozdělení, jež je součástí knihovny PyTorch (URL16). V souladu s výše uvedeným pak  $P(D_t|\theta)$  zastupuje vliv dílčích výstupů BNN na  $P(\theta|D_t)$  a  $P(D_t)$  je konstantní vůči parametrům sítě  $\theta$ , čímž normalizuje  $P(\theta|D_t)$  (Goan a Fookes, 2020).

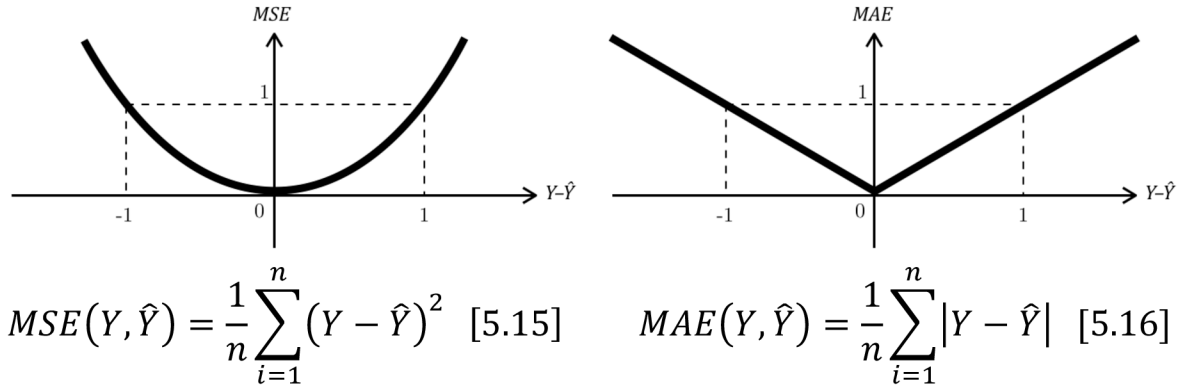
Pro tak komplexní modely, jakými jsou BNN, není reálně možné  $P(\theta|D_t)$  spočítat standardními metodami, jelikož určení integrálu při výpočtu  $P(D_t)$  je příliš obtížné (Jospin et al., 2022). Problém je ale možné obejít užitím rodiny sofistikovaných MCMC (*Markov Chain Monte Carlo*) algoritmů umožňující vzorkovat  $P(\theta|D_t)$  přímo nebo výpočetně rychleji užitím VI (*Variational Inference*) pro zjednodušenou aproximaci  $P(\theta|D_t)$  (Jospin et al., 2022). Detailnější rozbor zmíněných přístupů přesahuje možnosti této práce a poskytují je například Bishop (2006) a Goan a Fookes (2020).

Ne všechny vrstvy v BNN musí být stochasticky parametrizovány. Například Zeng et al. (2018) považují za efektivní kompromis umístění jedné nebo dvou bayesovských vrstev poblíž výstupu sítě, jelikož výpočetní náročnost plně stochastické sítě může být velmi vysoká.

### 5.1.7 Ztrátové funkce

Volba ztrátové funkce (anglicky *loss function*) je zásadní při práci s výstupem ANN a vychází z konkrétních potřeb dané aplikace (Aggarwal, 2018). Slouží k stanovení chyby ANN při vykonávání svého úkolu (Kruse et al., 2022) a tudíž hraje stěžejní roli při učení (viz podkapitola 5.1.8) (Goodfellow et al., 2016).

Autoregresní povaha modelu implikuje užití ztrátové funkce pro regresní problémy, z nichž MSE (*Mean Squared Error*) a MAE (*Mean Absolute Error*) patří k často používaným (Botchkarev, 2019). MSE je v této práci využito při učení ANN, jelikož díky své mocninné povaze vede k rychlejší konvergenci parametrů při učení, tj. velké chyby penalizuje přísněji a velmi malé chyby benevolentněji (Wang et al., 2022). Vlastností MAE je v této práci využito u mapových vizualizací chybovosti predikcí (viz podkapitola 6.3), jelikož zachovává stejné jednotky jako vstupní data a je tedy lépe interpretovatelná (Botchkarev, 2019). Na Obr. 17 jsou obě ztrátové funkce uvedeny včetně jejich funkčních předpisů:



Obr. 17 Ztrátové funkce MSE a MAE, upraveno z Wang et al. (2022) a Botchkarev (2019).

kde  $Y$  je tenzor skutečných hodnot, vůči nimž je hodnocen tenzor výstupů  $\hat{Y}$  generovaných sítí během učení (viz podkapitola 5.1.8) a  $n$  je počet prvků v tensech  $Y$  a  $\hat{Y}$ . Knihovna PyTorch umožňuje s chybami provádět i jiné operace, než je průměrování (URL17 nebo URL18).

Knihovna BLiTZ po vzoru Blundell et al. (2015) využívá VI k vytvoření parametrizovaného empirického rozdělení  $Q(\theta)$  (URL19), které vstupuje do evaluace  $P(\theta)$  (tj. očekávaného rozdělení parametrů sítě  $\theta$ ), užitím ztrátové funkce ELBO (*Evidence Lower Bound*) (Blundell et al., 2015). K jejímu užití je však nejprve potřeba představit princip Kullback-Leiblerovy divergence (dále KLD), kterou lze interpretovat jako míru neshody mezi dvěma rozděleními pravděpodobnosti (Bishop, 2006). Výpočet KLD pro spojitá rozdělení poskytuje Bishop (2006):

$$KLD(A \parallel B) = - \int A(x) \ln \left( \frac{B(x)}{A(x)} \right) dx \quad [5.17]$$

kde  $A$  a  $B$  jsou libovolná rozdělení pravděpodobnosti.

Blundell et al. (2015) představují výpočet ELBO po dílčích úpravách jako:

$$ELBO(D_t, \eta) = KLD(Q(\theta|\eta) \parallel P(\theta)) - \mathbb{E}_{Q(\theta|\eta)}(\log P(D_t|\theta)) \quad [5.18]$$

kde  $\eta$  jsou parametry metodou VI získaného aproximovaného rozdělení  $Q(\theta)$  a  $\mathbb{E}_{Q(\theta|\eta)}(\log P(D_t|\theta))$  je očekávaná hodnota tzv. *log-likelihood*, které obdobně jako samotný *likelihood*  $P(D_t|\theta)$  reprezentuje vliv modelu.

Výpočet lze interpretovat tak, že KLD je závislé na  $P(\theta)$  a  $\mathbb{E}_{Q(\theta|\eta)}(\log P(D_t|\theta))$  na  $D_t$ , což reprezentuje kompromis mezi potřebou reflektovat komplexitu dat  $D_t$  a relativní jednoduchostí  $P(\theta)$  (Blundell et al., 2015).

Knihovna BLiTZ používá také označení ELBO, nicméně výpočet mírně modifikuje. Datově závislou složku  $\mathbb{E}_{Q(\theta|\eta)}(\log P(D_t|\theta))$  nahrazuje kritériem jiné (např. regresní) ztrátové funkce, která také explicitně popisuje schopnost sítě modelovat data  $D_t$ . Transkripční kódu z URL20 do matematického zápisu je získán výpočet celkové chyby jako:

$$E = \frac{1}{k} \sum_{i=1}^k \beta \text{KLD}(Q(\theta|\eta) \parallel P(\theta)) + \text{MSE}(Y, \hat{Y}_i) \mid \hat{Y}_i = \Phi(X, \theta_i) \quad [5.19]$$

kde  $E$  je hodnota ELBO pro konkrétní průchod BNN psané v BLiTZ,  $\beta$  je váha hodnot KLD,  $\Phi$  zastupuje BNN vnímanou jako funkci (viz podkapitola 5.1.8) pro vstup  $X$  s parametry  $\theta_1, \dots, \theta_k$  získanými výběrem  $k$  vzorků parametrů  $\theta$ , pro něž jsou spočítány výstupy sítě  $\hat{Y}_1, \dots, \hat{Y}_k$ . Samotný výpočet KLD je z praktických důvodů v BLiTZ také mírně modifikován (URL21).

V této práci zvolené MSE může být nahrazeno i jinou ztrátovou funkcí podle účelu BNN, což umožňuje větší flexibilitu při řešení rozmanitých úloh.

### 5.1.8 Algoritmus učení

Prvním krokem trénování ANN je rozdělení původní množiny datasetu  $D$  na trénovací množinu  $D_t \subseteq D$ , na níž je ANN iteračně trénována, a validační množinu  $D_v \subseteq D$ , pomocí které je ověřována schopnost ANN vytvářet přesné výstupy i pro nové vstupy, přičemž platí, že  $D_t \cap D_v = \emptyset$  (Kruse et al., 2022).

Pro potřeby této práce je využito tzv. učení s učitelem (anglicky *supervised learning*), charakterizovaného učením asociací mezi vstupy  $X$  a výstupy  $Y$  obsaženými v  $D_t$  (Goodfellow et al., 2016). Ty jsou organizovány do podoby  $p$  uspořádaných 2-tic tak, že  $D_t = \{(X_1, Y_1), \dots, (X_p, Y_p)\}$  (Jindal et al., 2020). Mechanismy tzv. *batch learning*, přestože jsou mnohdy užitečné, nejsou v této práci aplikovány, a tudíž ani rozebrány.

Pokud jsou výsledky pro  $D_t$  výrazně přesnější než pro  $D_v$ , dochází k fenoménu tzv. přeučení, kdy síť ztrácí potřebnou schopnost generalizace (Aggarwal, 2018). Kromě výše uvedených metod lze přeučení bránit také předčasným ukončením učení, užitím různých regularizací při učení a snížením počtu neuronů v síti (Kruse et al., 2022).

V návaznosti na poznatky z podkapitoly 5.1.3 lze výstup generovaný ANN  $\hat{Y}$  vnímat jako výstup složitějšího řetězce vnořených funkcí, kdy každá vrstva je funkcí výstupu předcházející vrstvy (resp. u vstupní vrstvy vstupů  $X$ ) a svých parametrů  $\theta$  (Goodfellow et al., 2016). Zápis této skutečnosti poskytuje Aggarwal (2018), což lze na příkladu 3-vrstvé ANN modifikovat jako:

$$\hat{Y} = \varphi_3 \left( L_3 \left( \varphi_2 \left( L_2 \left( \varphi_1 \left( L_1(X, \theta_1) \right), \theta_2 \right), \theta_3 \right) \right) \right) \sim \Phi(X, \theta_1, \theta_2, \theta_3) \quad [5.20]$$

kde  $L_1, L_2, L_3$  jsou dílčí vrstvy ANN s parametry  $\theta_1, \theta_2, \theta_3$  a  $\varphi_1, \varphi_2, \varphi_3$  jsou k nim náležící aktivační funkce. Zástupná funkce  $\Phi$  pak umožňuje celý výpočet (resp. ANN) zapsat stručněji.

Chybu pro  $l$ -vrstvou ANN je pak možné v návaznosti na Aggarwal (2018) spočítat jako:

$$E = \varepsilon(Y, \Phi(X, \theta_1, \dots, \theta_l)) \quad [5.21]$$

kde  $E$  je chyba pro 2-tici  $\langle X, Y \rangle$  při učení sítě za užití ztrátové funkce  $\varepsilon$  a parametrů  $\theta_1, \dots, \theta_l$ .

Učení ANN je typicky prováděno algoritmem *backpropagation* (dále BP) (Aggarwal, 2018). V jádru algoritmu BP je optimalizační problém hledající takové  $\theta_1, \dots, \theta_l$ , které minimalizují  $\varepsilon$ , k čemuž je použito heuristické metody tzv. gradientního sestupu (anglicky *gradient descent*, dále GD) těžícího z poznatků diferenciálního počtu, především tzv. řetízkového pravidla (anglicky *chain rule*) pro vnořené funkce (Kruse et al., 2022). Proto musí být všechny  $\varphi_1, \dots, \varphi_l$  diferencovatelné (Goodfellow et al., 2016). Aplikaci tohoto pravidla na ANN poskytuje Haykin (2009) a v návaznosti na příklad v rovnici 5.20 by mohla mít s cílem derivovat například parametry  $\theta_3$  tvar:

$$\frac{\partial \varepsilon}{\partial \theta_3} = \frac{\partial \varepsilon}{\partial \varphi_3} \cdot \frac{\partial \varphi_3}{\partial L_3} \cdot \frac{\partial L_3}{\partial \theta_3} \quad [5.22]$$

Těchto vlastností je využito při určování tzv. gradientů (tj. tenzorů parciálních derivací pro každý parametr každé vrstvy) v jehož záporném směru jsou následně hodnoty všech parametrů  $\theta_1, \dots, \theta_l$  iterativně aktualizovány pro stanovený počet epoch (Kruse et al., 2022). Samotné gradienty parametrů a proces jejich aktualizace během učení jsou pro základní variantu GD v návaznosti na Haykin (2009) popsány na tenzoru vah  $W$  jako:

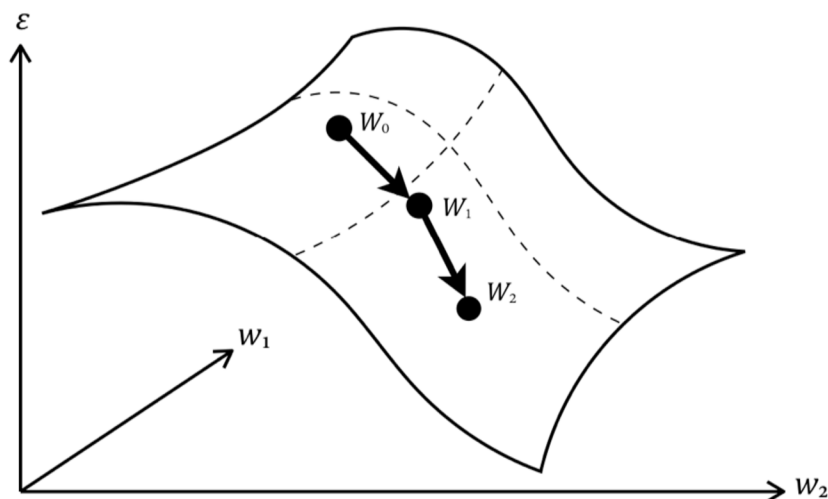
$$\nabla \varepsilon(W_{e-1}) = \begin{bmatrix} \frac{\partial \varepsilon}{\partial w_{11}} & \frac{\partial \varepsilon}{\partial w_{12}} & \dots & \frac{\partial \varepsilon}{\partial w_{1n}} \\ \frac{\partial \varepsilon}{\partial w_{21}} & \frac{\partial \varepsilon}{\partial w_{22}} & \dots & \frac{\partial \varepsilon}{\partial w_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \varepsilon}{\partial w_{m1}} & \frac{\partial \varepsilon}{\partial w_{m2}} & \dots & \frac{\partial \varepsilon}{\partial w_{mn}} \end{bmatrix} \quad [5.23]$$

$$W_e = W_{e-1} - \lambda \nabla \varepsilon(W_{e-1}) \quad [5.24]$$

kde  $\nabla$  označuje gradient,  $e$  označuje určitou epochu během učení a  $\lambda$  je tzv. koeficient učení (anglicky *learning rate*) regulující prudkost učení sítě, která může být v průběhu učení různě modifikována (Aggarwal, 2018), což v knihovně PyTorch zajišťují různé typy tzv. *schedulers* (URL22). Podobné výpočty lze aplikovat i na ostatní parametry ANN.

Ve výsledku lze princip GD vizualizovat způsobem uvedeným na Obr. 18, přičemž počet dimenzí optimalizačního prostoru (níže 3) odpovídá počtu všech parametrů + 1 pro hodnotu  $\varepsilon$ , což u reálných ANN může dosahovat řádově až miliard (Goodfellow et al., 2018).





Obr. 18 Princip GD pro příklad 2 epoch učení vah  $w_1$  a  $w_2$ , upraveno z Kruse et al. (2022).

Jelikož GD s sebou nese velké riziko spočinutí v nevhodném lokálním minimu, byly vyvinuty jeho robustnější alternativy, které obvykle implementují stochastické principy, obsahují výpočetní analogii k fyzikálnímu jevu setrvačnosti a/nebo pracují s různými normalizacemi gradientů (Kruse et al., 2022). Známými zástupci jsou např. SGD (*Stochastic GD*), RMSProp, AdaDelta, AdaGrad a v této práci použitý Adam (Kruse et al., 2022), který spolu se svými výhodami a potřebnými důkazy představili Kingma a Ba (2015).

Pro RNN bývá využita rozšířená varianta BPTT (*BP Through Time*) umožňující do procesu učení efektivně zakomponovat všechny dílčí interně sdílené tenzory vah spolu s délkou vstupní časové řady (Aggarwal, 2018). Algoritmus BPTT popisují například Goodfellow et al. (2016).

Pro BNN lze využít algoritmus *Bayes by backprop* od Blundell et al. (2015) umožňující poměrně přímočaré napojení na již existující postupy a rozličné výše zmíněné optimalizační přístupy (Jospin et al., 2022). V principu dochází k velmi podobnému výpočtu jako u deterministických sítí s tím rozdílem, že namísto konkrétních konstant je učení aplikováno na parametry pravděpodobnostních rozdělení parametrů sítě (Blundell et al., 2015).

### 5.1.9 Automatická optimalizace hyperparametrů

Pod pojmem hyperparametry jsou myšleny předem stanovené hodnoty v procesu designu ML algoritmu, přičemž v případě ANN jich je obvykle velké množství (Aggarwal, 2018). Hyperparametry nejsou narozdíl od parametrů nijak aktualizovány algoritmem učení, je však možné sestavit dva vnořené ML algoritmy tak, že jeden hledá nejlepší hyperparametry pro ten druhý (Goodfellow et al., 2016).

Společně s prudkým rozvojem AI roste poptávka po obecných nástrojích a inteligentních metodách se snahou snížit lidský faktor při řešení nových problémů, což je dnes pro mnoho aplikačních oblastí řešeno automatizací samotného procesu designu ML algoritmu (Qu, 2021). Za hlavní motivaci lze považovat skutečnost, že účinnost mnoha dnešních ANN je velmi citlivá na různé subjektivní volby programátora, jako je např. architektura ANN a její hyperparametry, volba učícího algoritmu, regularizační metody atd., což výrazně zvyšuje potřebu dlouhého testování různých variant modelů (Feurer a Hutter, 2019). Metody automatizovaného ML naopak umožňují snížení finančních a časových nároků, vyšší přístupnost ML technologií vědcům v rozličných aplikačních



oblastech a vyšší objektivitu výsledného modelu, navíc v mnoha ohledech poskytují lepší výsledky, než lidští ML experti (Hutter et al., 2019).

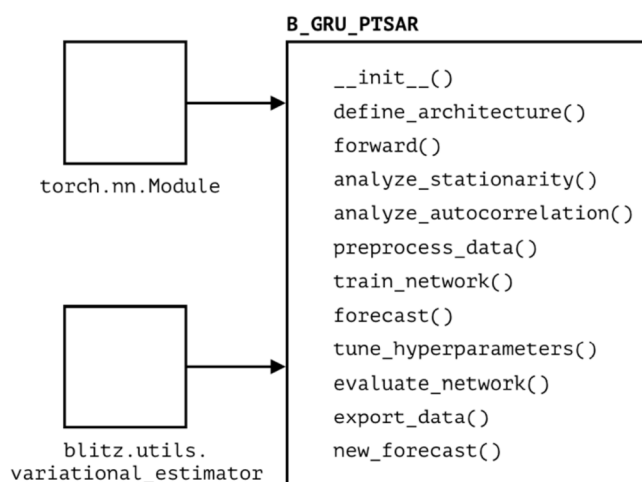
Mezi používané strategie optimalizace hyperparametrů patří např. *grid search*, *random search*, tzv. populační metody (např. optimalizace hejnem částic, evoluční a genetické algoritmy atd.) a bayesovská optimalizace (Waring et al., 2020). Knihovna Optuna poskytuje široké spektrum algoritmů, z nichž má každý své výhody a nevýhody (viz URL23).

Pro tuto práci je relevantní algoritmus CMA-ES (*Covariance Matrix Adaptation Evolution Strategy*), jeden z nejznámějších představitelů populačních metod provádějící výběr hyperparametrů z vícerozměrného Gaussova rozdělení, jehož průměr a kovariance je v duchu biologické evoluce iterativně aktualizována v generacích, přičemž do následující generace postupují pouze nejlepší řešení, která se mezi sebou následně různě kombinují a případně náhodně mutují (Feurer a Hutter, 2019).

Problematika evolučních algoritmů přesahuje možnosti této práce a věnuje se jim například Zhou et al. (2019). Stejně tak samotný CMA-ES algoritmus není pro svou komplexnost dále popisován a jeho rozbor poskytuje Hansen (2016).

## 5.2 Třída B-GRU-PTSAR

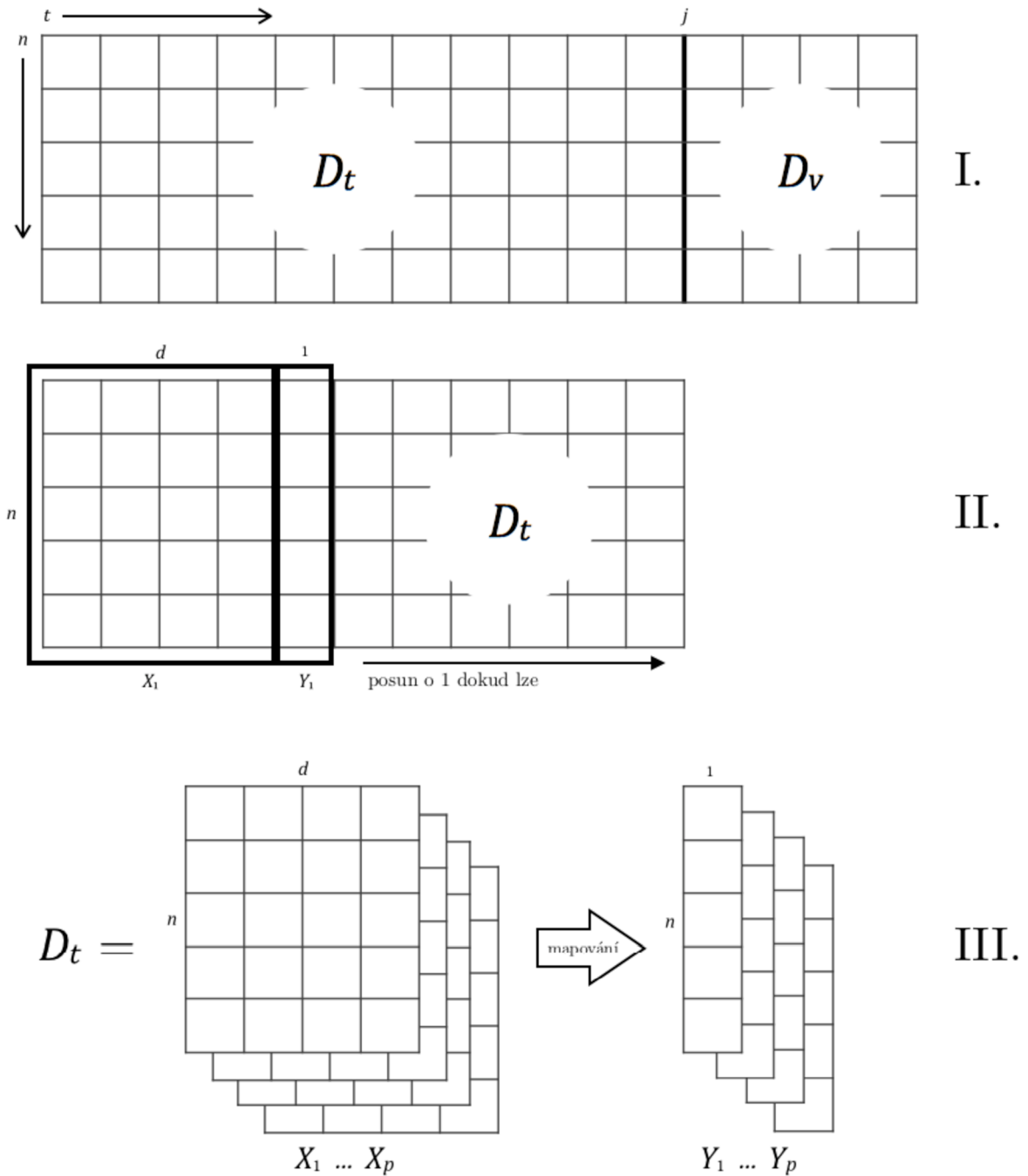
Za účelem tvorby ANN pro potřeby predikování prevalence COVID-19 v ORP ČR byla autorem práce sestavena obecná třída B-GRU-PTSAR (*Bayesian GRU Parallel Time Series Autoregressor*). Výhodou tohoto přístupu je kromě snadné přenositelnosti kódu do jiných skriptů také možnost jednoduše vytvářet vlastní ANN pro podobné úlohy zahrnující predikce paralelních časových řad. Při tvorbě bylo využito principů OOP (objektově orientované programování), jimž se věnují například Dathan a Ramnath (2015). Třída aplikuje teorii ze 4. kapitoly, z podkapitoly 5.1 a zároveň umožňuje vyhodnocení modelu z 6. kapitoly. Schéma třídy zobrazuje Obr. 19.



Obr. 19 Schéma třídy B-GRU-PTSAR a k ní stanovená dědičnost.

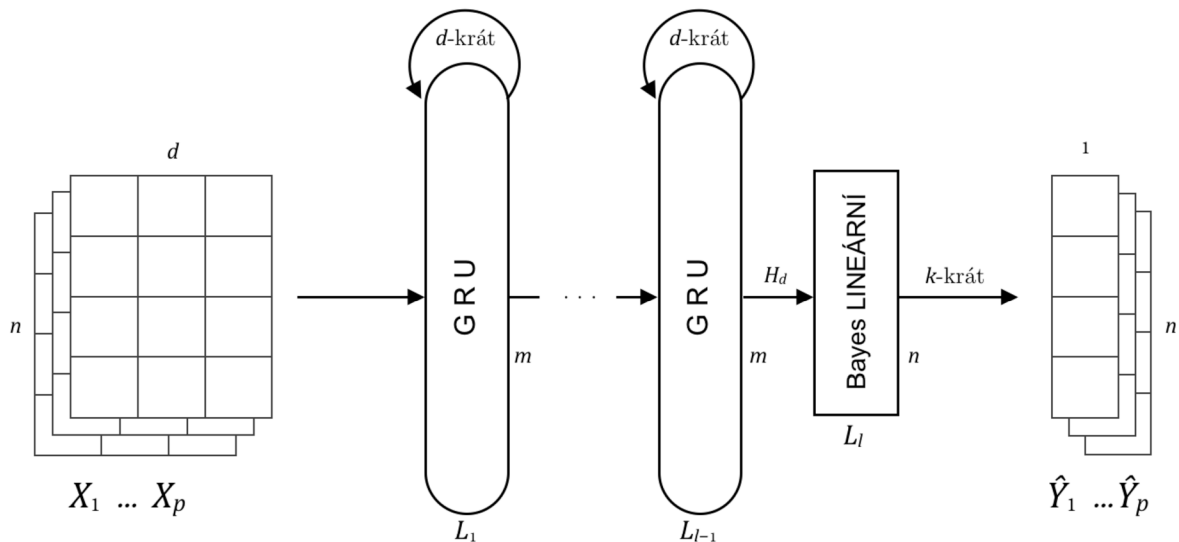
Třída předpokládá data v rozměru  $n \times t$  a typu `numpy.ndarray`, kde  $n$  je počet časových řad a  $t$  je jejich délka. Data jsou následně v bodě  $j$  rozdělena na trénovací množinu  $D_t$  a validační množinu  $D_v$ . Množina dat  $D_t$  je následně zpracována do tenzorového datasetu metodou tzv. posuvného okna, kdy je mírou autokorelace určená výšeč minulosti o délce  $d$  mapována na následující časový okamžik, čímž vzniká množina  $p$  uspořádaných

2-tic. Množina dat  $D_v$  slouží k validaci přesnosti predikcí mimo trénovací data po dokončení učení. Schéma procesu zobrazuje Obr. 20.



Obr. 20 Schéma procesu vzniku tenzorového datasetu pro učení.

Po rozsáhlém empirickém testování a srovnávání různých variant byla zvolena za třídou generovanou architekturu  $l$ -vrstvá BNN s  $l - 1$  deterministickými *stacked* GRU vrstvami o šířce  $m$ , zakončená lineární bayesovskou vrstvou o šířce  $n$  provádějící  $k$  náhodných výběrů svých parametrů, díky čemuž obsahuje výstup sítě informaci o nejistotě predikcí. BNN je připravena na příjem  $p$  dat o rozměru  $n \times d$  a po průchodu sítě má výstup rozměr  $n \times 1$ . Schéma této architektury zobrazuje Obr. 21.



Obr. 21 Schéma architektury sítě generované třídou B-GRU-PTSAR.

Před využitím třídy je potřeba surová data importovat a restrukturalizovat do vhodné podoby (pro data zmíněná v podkapitole 3.3 užitím knihovny Pandas viz Příloha 2). Tento proces není automatizovaný, jelikož různá surová data mohou mít různou strukturu a vyžadují odlišné operace. Třída obsahuje metody:

- `__init__()` – volána implicitně spolu s vytvořením instance třídy, ukládá vstupní data způsobem umožňujícím jejich další zpracování a inicializuje 0. řád difference,
- `define_architecture()` – slouží k definování šířky a hloubky generované sítě a umožňuje definovat *dropout* pro GRU vrstvy,
- `forward()` – definuje průchod sítě a vrací výstup výstupní vrstvy, takže slouží především jako nástroj pro další metody třídy, které pro svůj chod potřebují produkovat koncový výstup sítě,
- `analyze_stationarity()` – aplikuje ADF test postupně na všechny časové řady s tím, že je-li alespoň jedna řada nestacionární, lze na data aplikovat diferenci a test rekurentně opakovat do doby, než jej splní všechny řady,
- `analyze_autocorrelation()` – vykresluje graf ACF pro průměr všech časových řad, což lze použít při určení vhodné velikosti posuvného okna,
- `preprocess_data()` – dělí data na trénovací množinu  $D_t$  a validační množinu  $D_v$ , provádí také zvolenou transformaci  $D_t$  (lze vybrat jednu ze tří transformací probraných v podkapitole 4.3) a vytváří tenzorový dataset,
- `train_network()` – zahrnuje celý učicí cyklus sítě, tj. iterativní průchod  $D_t$  sítě a aktualizaci parametrů učícím algoritmem Adam, vyhodnocuje také dostupnost kompatibilního GPU pro účely hardwarové akcelerace učení a pokud jej nenalezne, zůstane výpočty na CPU (*Central Processing Unit*), dále umožňuje využít *scheduler* `torch.optim.lr_scheduler.ExponentialLR`, pomocí něhož lze během epoch učení exponenciálně snižovat hodnotu koeficientu učení  $\lambda$  během učení podle vzorce z URL24 po stanovení konstanty  $\gamma$  uživatelem jako:

$$\lambda_e = \gamma \lambda_{e-1} \quad [5.25]$$

- `forecast()` – generuje predikce pro stanovený počet dnů od konce  $D_t$ , které jsou vnitřně uloženy a později dále vyhodnoceny, přičemž mechanismus posuvného okna je zachován a funguje tak, že sítí projde posledních  $d$  záznamů, k nimž přiřazená nejpravděpodobnější predikce (průměr  $k$  výstupů) je připojena na konec kopie  $D_t$  a celý proces se opakuje, čímž jsou časové řady extrapolovány sítí stejným principem, jakým se je učila,
- `tune_hyperparameters()` – poskytuje nástroj pro automatizovanou optimalizaci hyperparametrů sítě algoritmem CMA-ES, což je vhodné k exploratorním účelům, ale také ke zcela automatizovanému sestavení finální podoby neuronové sítě, hyperparametry je možné fixovat na určité hodnotě a tím je z optimalizace vyřadit, s tím, že hlavním kritériem pro evaluaci vhodnosti hyperparametrů je přesnost predikcí a po vykreslení grafů na konci metody lze k dalšímu využití načíst ten nejlepší (ale i ostatní) příkazem `torch.load()` (viz URL25),
- `evaluate_network()` – slouží k vytvoření všech grafů potřebných k evaluaci procesu učení a přesnosti predikcí, předtím provádí inverzní transformaci a dediferenci predikcí (pokud byla data diferencována), což umožňuje jejich srovnání s  $D_v$ ,
- `export_data()` – uzavírá trénování třídy B-GRU-PTSAR generováním finálních výstupních tabulek ve formátu CSV (*Comma Separated Value*), které lze následně importovat do prostředí GIS (geografický informační systém) a propojit pomocí zvoleného primárního klíče s atributovou tabulkou odpovídajících prostorových dat,
- `new_forecast()` – umožňuje vygenerovat predikce užitím již natrénovaného modelu, je ale třeba mít na paměti, že vstupní data metody (např. aktualizovaný počet aktuálně nakažených) musí nejprve projít všemi diferencemi zvlášť, výstupem metody je nakonec matice obsahující všech  $k$  dílčích výstupů pro všech  $n$  objektů (např. ORP) po stanovenou dobu, kterou je potřeba dediferencovat s tím, že tato surovější forma byla zvolena z důvodu maximální flexibility následného využití.

Přesná struktura výše stručně charakterizovaných metod je dostupná ze zdrojového kódu v *Příloha 1*. Další informace lze nalézt v dokumentaci metod třídy B-GRU-PTSAR v *Příloha 3*.

## 6 VALIDACE MODELU

Na základě ADF testu bylo 104 z 206 vstupních časových řad nestacionárních. To však nemělo na přesnost predikcí žádný vliv a jejich diferencováním naopak stoupala náročnost jejich učení sítí a klesala přesnost predikcí. Vykreslením ACF pro průměr časových řad bylo následně určeno za optimální velikost posuvného okna 80 dnů. Rozdělení na trénovací množinu  $D_t$  a validační množinu  $D_v$  bylo provedeno 30 dní od konce, což je také doba, pro kterou byla provedena predikce po skončení učení sítě k validaci přesnosti predikcí. Před vstupem do modelu byla na data aplikována Yeo-Johnsonova transformace, která měla také mírný efekt na stacionaritu časových řad, jelikož po jejím užití bylo nestacionárních 72 z 206.

Níže validovanému výslednému modelu odpovídá *Příloha 5*.

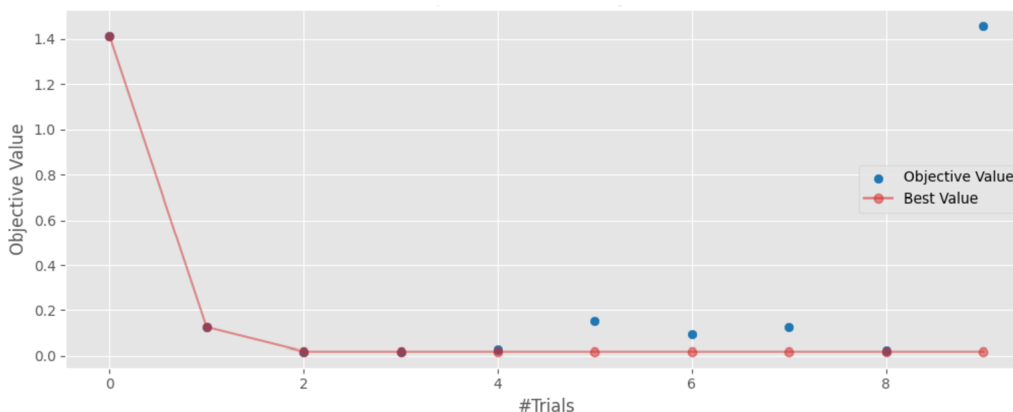
### 6.1 Validace automatické optimalizace hyperparametrů

Sestavení co nejlepšího modelu bylo s cílem maximalizace objektivnosti finálního modelu provedeno automatizovaně algoritmem CMA-ES, který otestoval 10 variant sítě, každou pro 100 epoch učení, což trvalo přibližně 8,5 hodin. Na výstupní bayesovské vrstvě byly vzorkovány 3 sady parametrů. *Tabulka 2* obsahuje stanovené rozsahy testovaných optimalizovaných hyperparametrů a ty nejlepší získané ve zcela nezaokrouhlené podobě, jelikož i nepatrná změna může vést ke zcela odlišným výsledkům.

*Tabulka 2* Rozsahy optimalizovaných hyperparametrů a jejich nejlepší hodnoty.

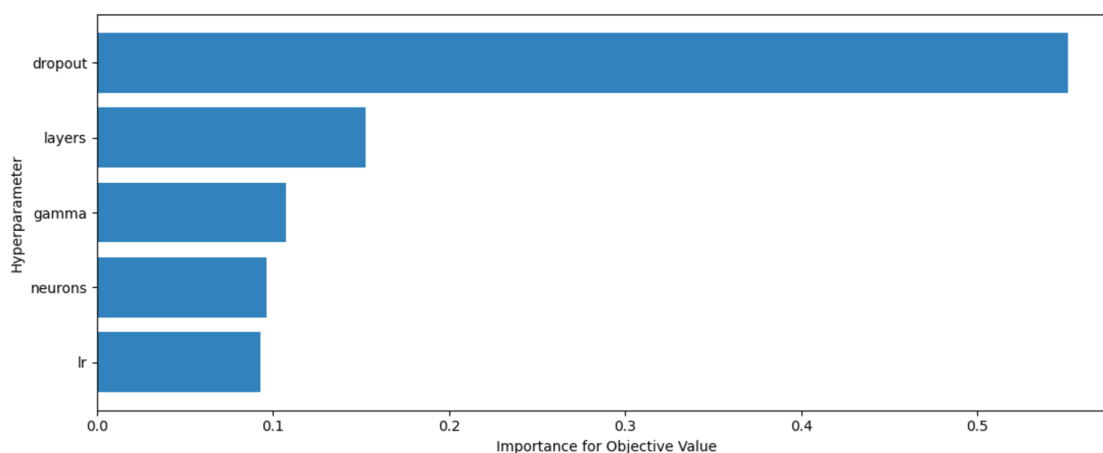
Hyperparametr	Stanovený rozsah	Nejlepší hodnota
koeficient učení	0,00001–0,001	0,0005925903420950085
počet neuronů ( $m$ )	200–800	500
dropout	0–0,3	0,1500614840529979
počet GRU vrstev ( $l - 1$ )	2–5	4
konstanta $\gamma$	0,9–1	0,9501072672133868

Knihovna Optuna umožňuje snadno graficky vizualizovat průběh celého procesu, zobrazeného na *Obr. 22*, na základě něhož (a také průběžných výpisů během optimalizace) je určena za nejlepší 2. testovaná varianta hyperparametrů sítě.



*Obr. 22* Proces automatické optimalizace hyperparametrů.

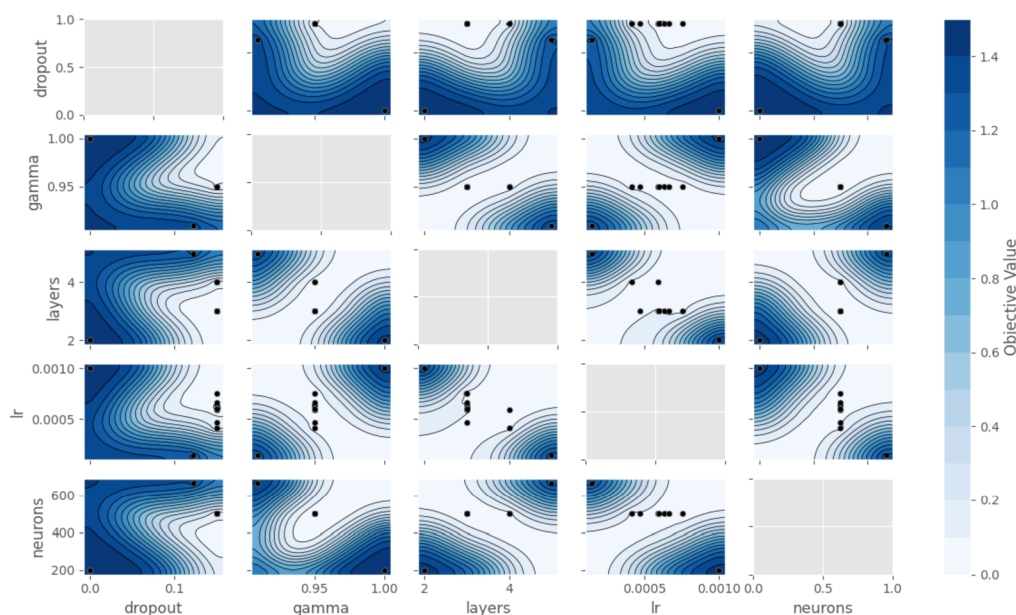
Další hodnotnou zakomponovanou funkcionalitou knihovny Optuna je možné vizualizovat významnost všech optimalizovaných hyperparametrů pro přesnost dané sítě, což je obsahem Obr. 23.



Obr. 23 Významnost hyperparametrů z hlediska přesnosti predikcí.

Výrazně vyšší významnost hodnoty *dropout* pro přesnost predikcí je poměrně překvapivá. Jedno z potenciálních vysvětlení může být nadměrná míra nahodilosti při deaktivaci velkého procenta neuronů, což v kombinaci se stochastickou povahou bayesovské vrstvy může vést ke vzniku velmi nestabilního modelu.

V knihovně Optuna je dále možné vizualizovat samotný optimalizační prostor hyperparametrů, jak je tomu na Obr. 24, díky čemuž je ve formě izoliniemi zobrazeného spojitého prostoru dále možné hodnotit jednotlivé hyperparametry. Kromě získání znalostí o vztazích mezi hyperparametry je tato informace využitelná při stanovování nových rozsahů testovaných hodnot v případných dalších experimentech.



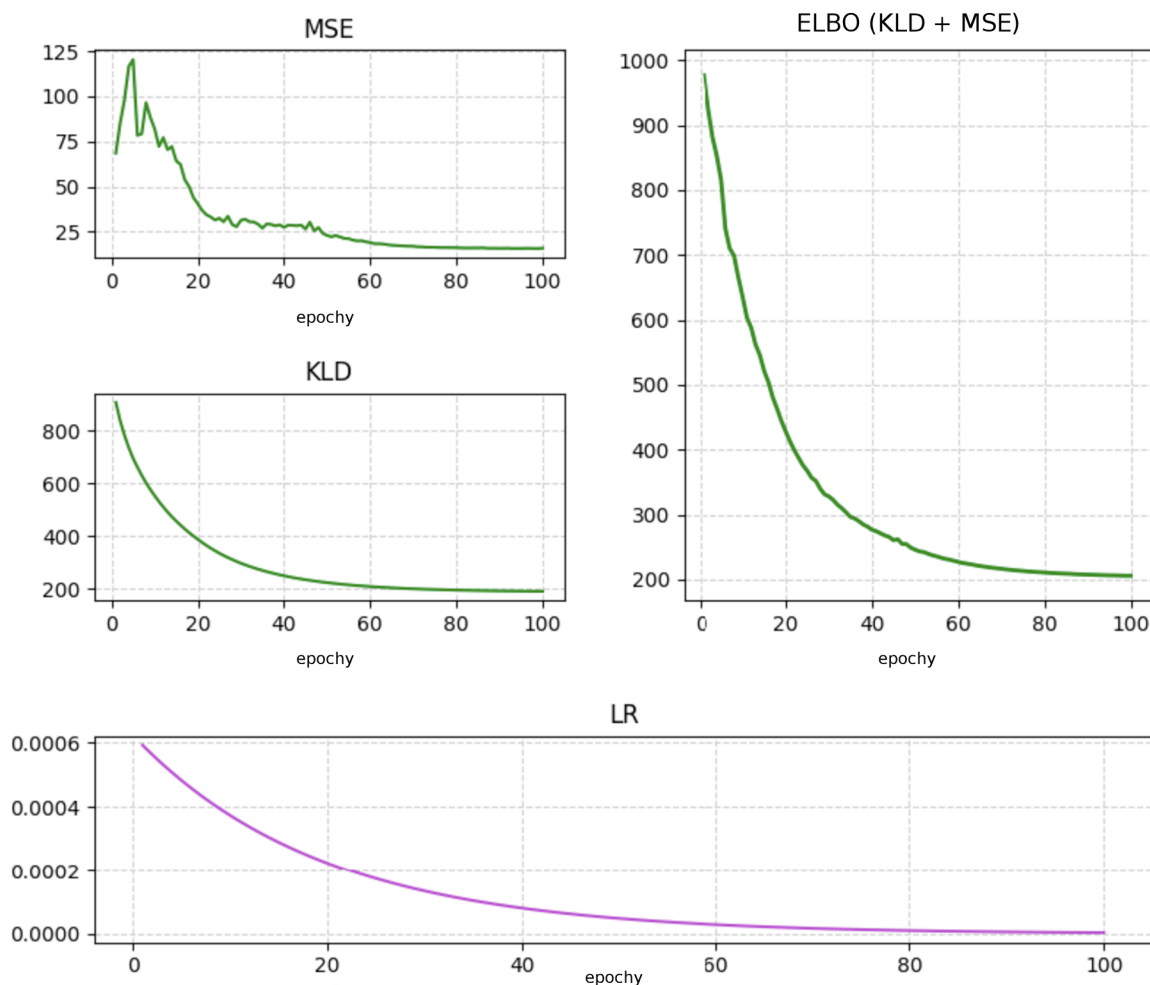
Obr. 24 Optimalizační prostor hyperparametrů.

(světlejší hodnotě odpovídá nižší hodnota chyby predikcí, tj. vyšší přesnost)

Na základě výše uvedených závěrů lze s vysokou mírou jistoty tvrdit, že všechny z testovaných rozsahů hyperparametrů byly vhodně zvoleny, jelikož algoritmus konvergoval v rámci jimi stanoveného optimalizačního prostoru.

## 6.2 Validace učení

Důležitým nástrojem pro získání vhledu do procesu učení je tzv. křivka učení (anglicky *learning curve*), přičemž u ANN v případě validního učení typicky nabývá přibližného tvaru klesající exponenciální křivky (Viering a Loog, In print). To splňuje i výsledný model, jak dokazují grafy na Obr. 25.

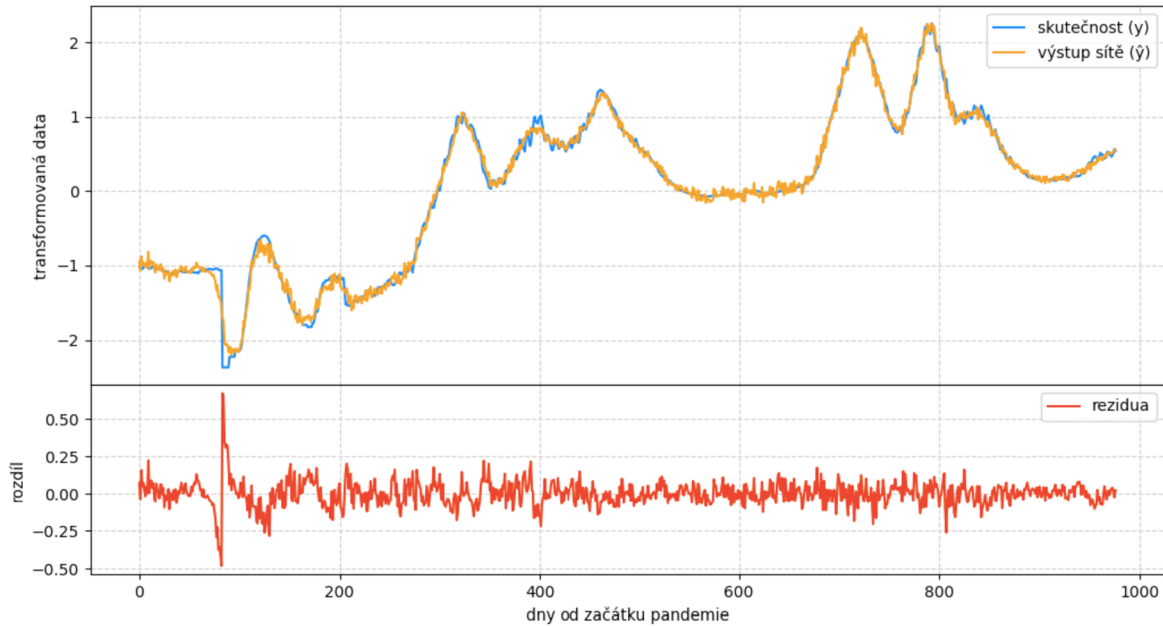


Obr. 25 Grafy procesu učení, resp. chování ztrátové funkce a koeficientu učení (zde značeno jako LR) napříč epochami učení.

Na základě grafů je zřejmé, že hodnota ztrátové funkce ELBO (tj. sumy KLD a MSE) v průběhu epoch klesala podle očekávání a asymptoticky se blížila přibližně k hodnotě 200. Při pohledu pouze na MSE část ELBO je zajímavá oscilace přibližně do 20. epochy, která může být způsobena zpočátku velmi nejasnými pravděpodobnostními rozděleními parametrů bayesovské výstupní vrstvy, přičemž stabilizace nastala až po jejich dostatečném zpřesnění. Méně významné oscilace v pozdějších epochách mohou pramenit také z klesající hodnoty koeficientu učení implikující méně výrazné aktualizace parametrů ke konci učení.



Síť se i přes výrazné stochastické prvky dokázala velmi přesně naučit trénovací množinu  $D_t$ , jak dokládá graf na Obr. 26 obsahující skutečné a modelem vygenerované hodnoty dílčích prvků  $D_t$  na příkladu ORP Olomouc po 100. epoše.



Obr. 26 Graf ztvárňující výkon modelu na trénovací množině  $D_t$ .  
(ORP Olomouc)

Tato skutečnost poukazuje na kapacitu sítě osvojit si velmi přesně modelovaná data, a to ve všech stavech procesu, tj. pro stoupání, klesání, konstantní průběh apod. Mírnou oscilací modelem rekonstruované trénovací množiny  $D_t$  lze považovat za přirozený důsledek stochasticity modelu.

Fenoménu přeučení bylo zamezeno užitím *dropout*, bayesovskou parametrizací výstupní vrstvy a kritériem automatické optimalizace hyperparametrů, kterým byla přesnost predikcí, tedy přesnost učení byla v tomto ohledu zcela irrelevantní.

Na základě výše zmíněných tvrzení a vizualizací lze s velkou mírou jistoty proces učení považovat za úspěšný a validní.

### 6.3 Validace predikcí

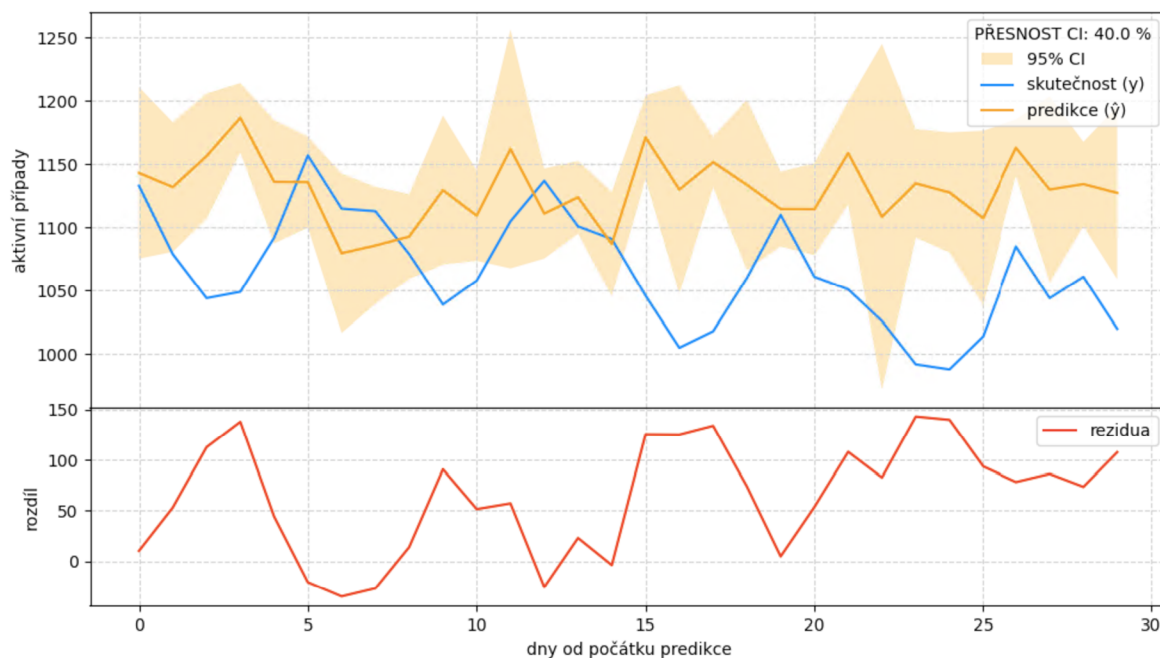
Na rozdíl od validace učení, kde stačilo pracovat s průměrem  $k$  náhodně vzorkovaných výstupů, jelikož z nich byla počítána hodnota ELBO, je u predikcí vhodné přidat informaci o jejich nejistotě. K tomu je využito intervalu spolehlivosti (anglicky *confidence interval*, dále CI), který v deskriptivní statistice slouží k vyjádření jistoty odhadu o stanovené hladině spolehlivosti  $\alpha$ , která bývá obvykle volena rovna 0,05 (Hazra, 2017). Výpočet CI obecně uvádí Hazra (2017) a v kontextu této práce je upraven na:

$$CI = \bar{y} \pm z \frac{\sigma}{\sqrt{k}} \mid \bar{y} = \frac{1}{k} \sum_{i=1}^k y_i, \sigma = \sqrt{\frac{\sum_{i=1}^k (y_i - \bar{y})^2}{k - 1}} \quad [6.1]$$



kde  $y_1, \dots, y_k$  jsou výstupy sítě vzniklé  $k$  vzorky parametrů bayesovské vrstvy,  $\bar{y}$  je jejich průměr,  $\sigma$  je výběrová směrodatná odchylka a  $z$  označuje tzv.  $z$ -score, které je pro  $\alpha = 0,05$  rovno 1,96 (Hazra, 2017). Výpočty  $\bar{y}$  a  $\sigma$  jsou doplněny z Wan et al. (2014).

Graf predikce pro konkrétní ORP pak vypadá jako na Obr. 27.

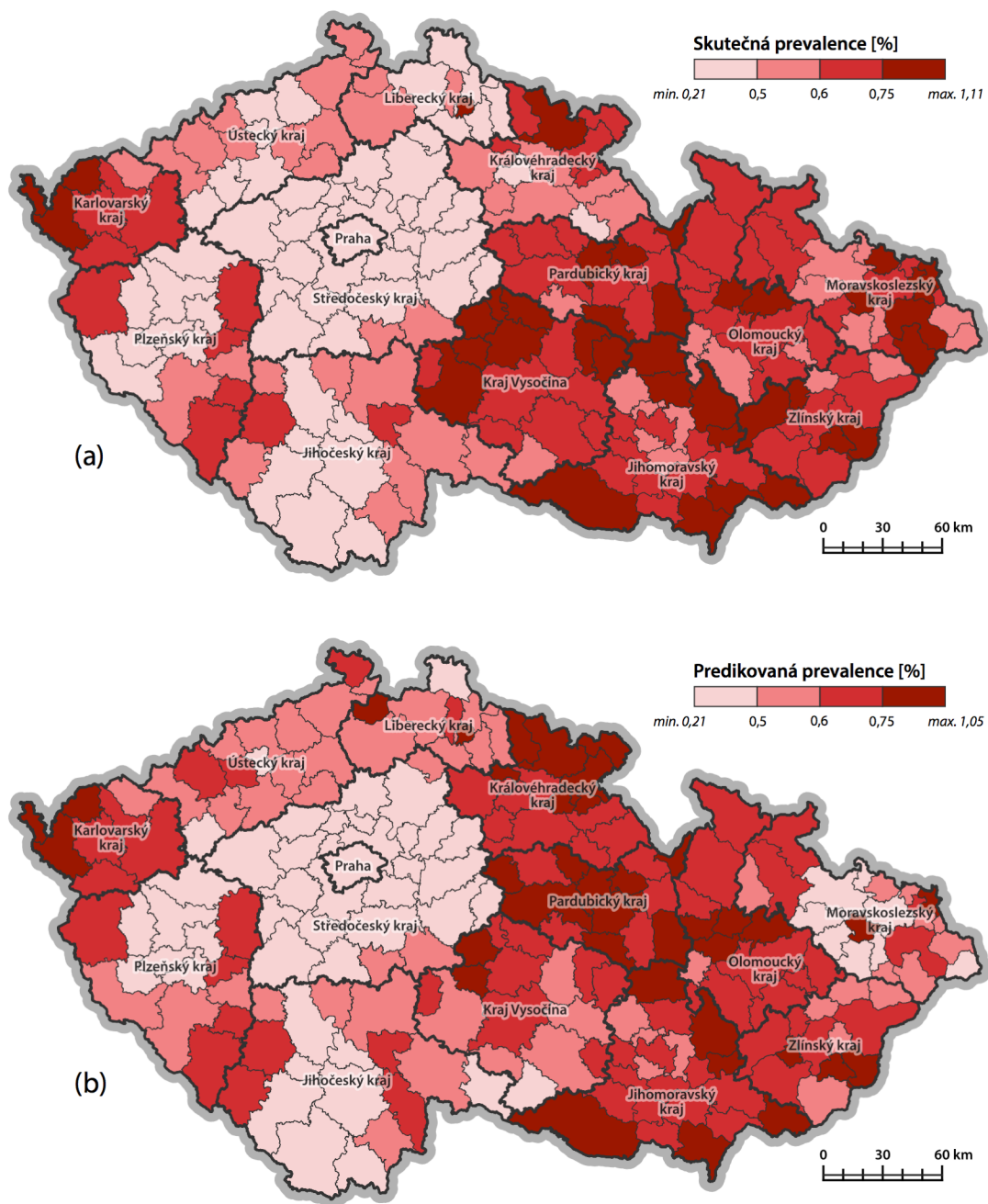


Obr. 27 Graf predikcí vůči validační množině  $D_v$ .  
(ORP Olomouc)

Pro konkrétní případ ORP Olomouc byl model schopen zachytit víceméně konstantní trend aktuálního počtu nakažených a chyba se za celé území pohybuje obvykle v řádu desítek osob, místy překračuje stovku. Stochastická povaha modelu dále poskytuje data o nejistotě predikcí ve formě CI, díky čemuž je možné informovaněji uvažovat o možném kolísání predikcí.

Přesnost CI označuje procento skutečných hodnot, které spadají do CI stanoveného na základě  $k$  dílčích výstupů. Potřeba takovéto charakteristiky pro validaci predikcí je logická, nicméně může být zavádějící, jelikož u velmi „nejisté“ sítě (tj. CI by byl velmi široký) by mohla snadno dosahovat až 100 %. Naopak situace, kdy si je model velmi jistý jen mírně chybnými predikcemi (tj. CI je kolem těchto hodnot velmi úzký), může vést k nízké přesnosti CI, přestože by se jinak jednalo o velmi přesné výsledky. Z tohoto důvodu je třeba charakteristiku interpretovat obezřetně a v kontextu s dalšími kritérii.

Jelikož síť pracuje současně se všemi ORP, je možné z dat v tabulkách vygenerovaných metodou `export_data()` vytvořit různé mapové vizualizace. Predikce i skutečný stav aktuálně nakažených lze snadno přepočítat na prevalenci tak, že jsou vyděleny počtem obyvatel z dat SLDB 2021, přičemž predikce byly nejprve celočíselně zaokrouhleny (tj. na celé osoby). Srovnání skutečné a predikované prevalence pro ukázkový 15. den predikce poskytují mapy na Obr. 28.

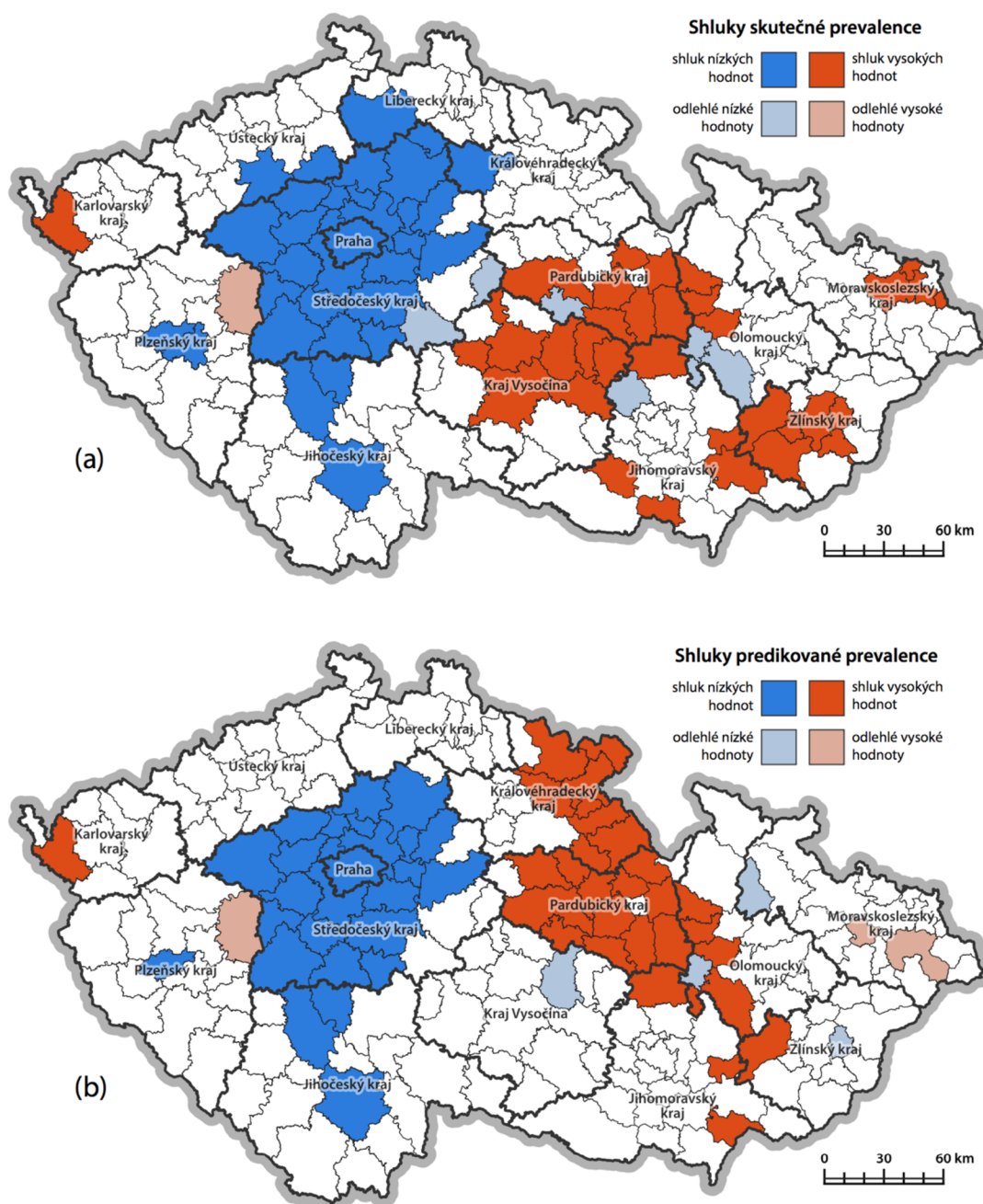


Obr. 28 Srovnání skutečné (a) a predikované (b) prevalence pro 15. den predikce.  
(ORP ČR + Praha)

Z výše uvedených map plyne schopnost modelu generovat predikce pro celé stanovené území, přičemž lze pozorovat podobnost skutečného a predikovaného stavu. Za zmínku stojí také velmi blízké minimální a maximální hodnoty viditelné na koncích stupnic obou map.

Predikce také na první pohled vytvářejí prostorové shluky podobně jako reálná data, což je možné robustněji ověřit užitím LISA (*Local Indicators of Spatial Autocorrelation*), tj. metodou prostorové statistiky identifikující odlehle hodnoty a prostorové shluky (Osadebey et al., 2019). Realizovat LISA je v prostředí QGIS snadné užitím pluginu Visualist, přičemž pro následující mapy bylo jako indikátor použito Moranovo I, sousedství typu královna

a hladina významnosti  $\alpha = 0,05$ . Mapy prostorových shluků pro skutečné a predikované hodnoty prevalence pro ukázkový 15. den predikce jsou zobrazeny na Obr. 29.

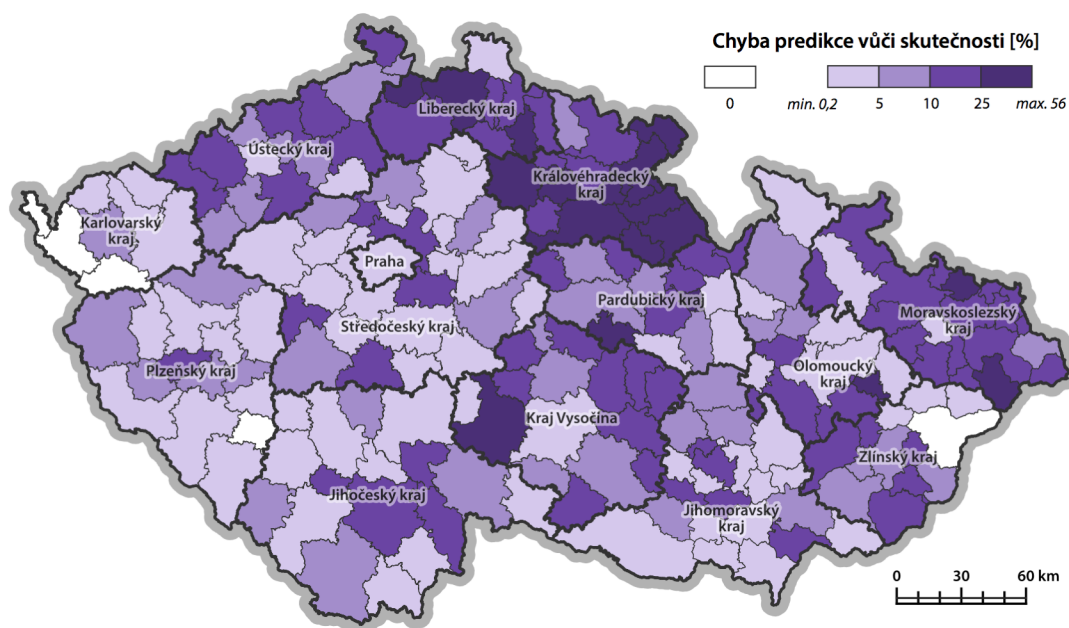


Obr. 29 Srovnání shluků skutečné (a) a predikované (b) prevalence pro 15. den predikce. (ORP ČR + Praha)

Predikce pro ukázkový 15. den vytvářejí shluky umístěné podobně jako skutečné hodnoty. Pro tento konkrétní den stojí za zmínku posunutí ohniska relativně vysoké prevalence severním směrem do Královéhradeckého kraje a podhodnocení prevalence v okolí ORP Bílovec (Moravskoslezský kraj). Shluk nízkých hodnot prevalence v oblasti Středočeského kraje a Prahy byl korektně identifikován, byť mírně menší. Za velmi

překvapivě lze považovat správné zařazení izolovaných území z hlediska shlukování, jako je např. ORP Cheb (Karlovarský kraj), ORP Rokycany a ORP Stod (Plzeňský kraj) apod. Zajímavé je také správné nezahrnutí ORP Česká Třebová (Pardubický kraj) do shluku vysokých hodnot prevalence.

Velmi hodnotné je vyhodnocení chybového kritéria, za něž bylo pro mapové vizualizace zvoleno MAE bez konečného průměrování, resp. absolutní hodnota rozdílu skutečných a predikovaných hodnot. Jelikož se chyba nechová stejně po celém území (např. v Ostravě je v řádu stovek, zatímco v méně zalidněných ORP je v řádu jednotek a desítek), byla relativizována na Obr. 30 jako poměr falešně pozitivních osob (po celočíselném zaokrouhlení) ku skutečnému počtu aktuálně nakažených osob pro ukázkový 15. den predikce.

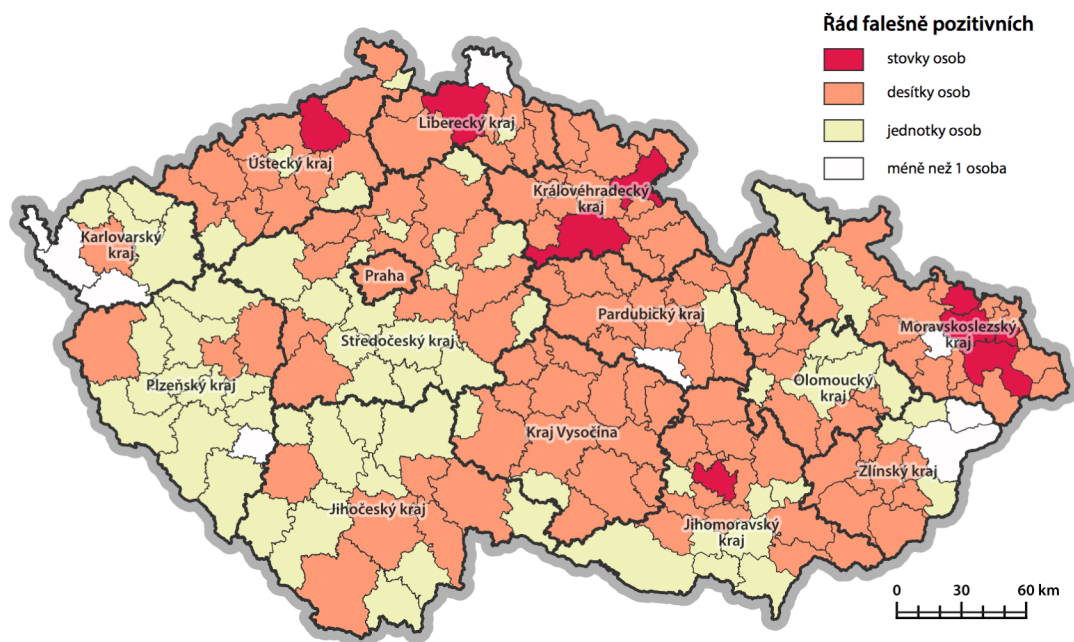


Obr. 30 Relativní chyba predikcí vůči skutečnosti pro 15. den predikce.  
(ORP ČR + Praha)

Vizualizace chyb reflektuje domněnku výraznější odchylky od skutečnosti v oblasti Královéhradeckého kraje a na východě Moravskoslezského kraje. V oblastech západních Čech byl model naopak velmi přesný. Za zmínku stojí lokálně výraznější chyba predikce pro ORP Stod (Plzeňský kraj), přestože daná oblast byla správně vyhodnocena jako oblast shlukování nízkých hodnot prevalence na předcházejícím Obr. 29. Tyto (ne)presnosti však nemusí být inherentně vlastní modelu a může se jednat o momentální výkyvy v dané fázi predikce, případně také mohou být způsobené specifickými vlastnostmi nedávného vývoje onemocnění COVID-19.

Kromě výše uvedené relativní chyby může být hodnotné kategorizovat jednotlivá ORP podle řádovosti absolutní hodnoty falešně pozitivních. To poskytuje mapa na Obr. 31.





Obr. 31 Řády falešně pozitivních osob pro 15. den predikce.  
(ORP ČR + Praha)

Z mapy vyplývá skutečnost, že pro převážnou většinu území ČR byla absolutní chyba modelem vytvořených predikcí v řádu jednotek a desítek osob. V několika ORP nacházejících se v oblasti méně přesných predikcí byla absolutní chyba v řádu stovek, jinde naopak i menší než 1 osoba.

Všechny výše uvedené mapy byly vztaheny k 15. dni predikce (tj. k polovině predikovaného období), je však samozřejmě možné na základě tabulek v Příloha 4 tvořit obdobné výstupy i pro další dny.

## 7 VÝSLEDKY

Celkového cíle práce bylo dosaženo naplněním dílčích cílů, tj. analýzou a předzpracováním vstupních dat, tvorbou ANN pro predikční účely a validací finalizovaného modelu.

Vybraná epidemiologická data počtu aktuálně nakažených od ÚZIS pokrývající území ČR (viz podkapitola 3.3) byla nejprve testována ADF testem za účelem potvrzení jejich stacionarity (viz podkapitola 4.1), kterým bylo 104 z 206 časových řad identifikováno jako nestacionární. Následná diference nebyla provedena na základě empirického testování. Vizualizací autokorelačních vazeb užitím ACF (viz podkapitola 4.2) byla určena velikost posuvného okna (viz podkapitola 5.2). Na data byla nakonec aplikována Yeo-Johnsonova mocninná transformace (viz podkapitola 4.3). Výsledkem dílčího cíle 4. kapitoly je tedy analyzovaný a předpřipravený dataset optimalizovaný pro potřeby predikčního modelování.

Krok samotné tvorby ANN je rozdělen na teoretičtější podkapitulu 5.1 představující všechny relevantní metody ANN a ML zvolené pro výslednou implementaci a praktičtější podkapitulu 5.2 stručně představující funkcionalitu autorem práce sestavené třídy B-GRU-PTSAR napsané v jazyce Python užitím zmíněných rozšíření (viz podkapitola 3.2), která byla použita pro tvorbu predikčního modelu nad epidemiologickými daty. Výsledkem dílčího cíle 5. kapitoly je tedy ANN pro predikční účely aktuálního počtu nakažených onemocněním COVID-19 na území ORP ČR a samotná třída B-GRU-PTSAR spravující tvorbu bayesovských RNN, ale i realizující předcházející analýzu a následnou validaci modelu.

Nejlepší vytvořený model byl následně validován zhodnocením průběhu automatické optimalizace hyperparametrů (viz podkapitola 6.1), učení ANN (viz podkapitola 6.2) a testováním přesnosti samotných predikcí (viz podkapitola 6.3). K tomu bylo kromě vybraných metrik využito několika datových a mapových vizualizací. Výsledkem dílčího cíle 6. kapitoly je tedy finální a v několika aspektech validovaný model.

Všechny dílčí výsledky jsou představeny v hlavní části textu a finální výstupy jsou součástí příloh.

V důsledku naplnění celkového cíle práce lze stanovit výsledné tvrzení, že je možné sestavit predikční model využívající metod ANN a ML pro účely predikčního modelování vybraných charakteristik pandemie COVID-19 na území ČR, a že tento model dosahuje uspokojivě přesných výsledků. Kritické zhodnocení práce poskytuje 8. kapitola.

## 8 DISKUZE

V kontextu již existujících přístupů (viz 2. kapitola) lze práci považovat za velmi moderní a v souladu se světovými trendy v oblasti (nejen) predikčního modelování infekčních onemocnění, kde se ANN a další ML přístupy těší intenzivnímu výzkumu a užití. V českém geoinformatickém prostředí představuje práce jeden z prvních pokusů tyto technologie představit širšímu okruhu odborné veřejnosti ve snaze rozvést diskuzi nad jejich využitelností v epidemiologii, ale i v jiných oblastech. K tomu je potřeba zaujmout kritické stanovisko.

Přestože jsou ANN a další ML metody obvykle považovány za jedny z nejefektivnějších, patří k jejich základní limitaci potřeba mnohdy rozsáhlých datasetů pro jejich natrénování. V rané fázi pandemie COVID-19, kdy by existence silného predikčního modelu mohla nejvýrazněji zmírnit dopad nemoci na společnost, by tedy pravděpodobně nebylo možné sestavit stejně spolehlivý model ANN, jako je možné dnes.

Komplikovaná je také samotná povaha dat, jelikož spolu s klesajícím zájmem veřejnosti o COVID-19 se snižuje počet provedených testů, tedy i počet identifikovaných případů. Za zmínku stojí také možnost samotestů, kdy soukromě identifikovanou pozitivitu nemusí nutně infikovaný jedinec nahlásit svému ošetřujícímu lékaři, v důsledku čehož se nepromítne do dat od ÚZIS. Z těchto (a pravděpodobně i dalších) důvodů je spolehlivost dat za poslední období sporná, navíc nevykazuje chování, které bylo viditelné na předchozích cyklech onemocnění, kdy bylo testování a vedení dat o COVID-19 bráno velmi vážně. Přesto je třeba konstatovat, že výsledný model dokázal při učení a v predikcích zachytit poměrně přesně i toto nečekané chování modelovaných dat.

Z hlediska analýzy a předzpracování dat se za problematické nabízí zachování nestacionarity v datech, přestože toto rozhodnutí bylo provedeno na základě empirického testování. Užití ACF pro průměr časových řad také není zcela optimální, jelikož korektně by každá z nich měla mít svůj vlastní rozbor autokorelace, aby nedocházelo k nadbytku dat nebo naopak ztrátě části informace. Nicméně i přesto lze představené využití ACF považovat za objektivnější a vhodnější, než by byla alternativa, tedy zkusmé hledání velikosti posuvného okna. Volba Yeo-Johnsonovy transformace ve finálním modelu vycházela z empirického testování, ale pravděpodobně by bylo možné najít argumenty i pro využití jiné transformace.

Ačkoliv všechny použité metody ANN byly adekvátně zvoleny vzhledem k řešenému problému, je jistě možné dosáhnout stejných (nebo lepších) výsledků i jiným přístupem. Celá oblast ML je natolik široká a rozmanitá, že možných řešení je skutečně mnoho, možná až nekonečno. Koneckonců i za zachování zde užitých metod je možné různě upravovat počet vrstev, jejich šířku, celkový tvar sítě, zvýšit počet bayesovských vrstev atd. Realita je taková, že i přes snahy zachovat maximální objektivitu při tvorbě modelu se jedná o výsledek mnoha subjektivních rozhodnutí.

Autorem vytvořená třída B-GRU-PTSAR splnila svůj účel pro naplnění cílů práce, její nároky na obecnější využití by však bylo nutné dále testovat. Z programátorského hlediska poskytuje dostatek možností na její optimalizaci, zpřehlednění kódu, rozšíření o další funkcionalitu a zlepšení uživatelských aspektů. Přestože její metody umožňují téměř vše nastavit při tvorbě vlastní ANN, některé náročnější úpravy mohou vyžadovat zásah přímo do jejího zdrojového kódu. Z těchto důvodů není vhodné třídu vnímat jako zcela hotový nástroj, ale spíše jako funkční prototyp.

Validaci modelu lze považovat za úspěšnou na základě zhodnocení několika kritérií, nebylo tak však učiněno vyčerpávajícím způsobem. Data vygenerovaná ve validační fázi modelování jsou poměrně početná a poskytnutí všech možných grafů a map pro

zhodnocení je tedy prakticky nemožné. Původním plánem bylo testovat model na různých fázích pandemie (nárůst, klesání apod.), nicméně kvůli výše popsaným problémům s daty od toho bylo upuštěno, jelikož by to vyžadovalo odstranění velké části již tak poměrně nepočetných dat z trénovací množiny. Snahy model co nejrobustněji validovat jsou také komplikovány značně uzavřenou povahou ANN. Je třeba také konstatovat, že jako každý jiný model bude i tento zatížen nedokonalostí a prostorem ke zlepšení. Koneckonců cílem nebylo představit definitivní řešení, ale pouze jeden z možných přístupů.

Praktické využití výsledného modelu je sice teoreticky možné, nicméně i tak by bylo potřeba provést dílčí úpravy v závislosti na konkrétním aplikačním využití a již zaběhlých systémech potenciálního uživatele. Za vhodné řešení lze považovat přesun modelu na webový server, odkud by denně přebíral aktualizovaná data, rovnou se z nich učil a výsledné predikce vykresloval formou interaktivních webových map. To v současném stavu modelu nemusí být možné. Realizace takového řešení nicméně přesahuje cíle i možnosti této práce, proto nebyla více prozkoumána.

Směrů navazujícího výzkumu je několik. Bylo by zajímavé sestavit komparativní studii pro mnoho různě sestavených ANN a porovnat jejich výsledky ve snaze nalézt nejlepší variantu pro predikování COVID-19 na území ČR. Dále, nebude-li shromažďování epidemiologických dat o COVID-19 zcela zrušeno, by také mohlo být zajímavé otestovat výkon modelu na dlouhodobé predikce, například půlroční. Hodnotných výsledků by mohlo být dosaženo obohacením autokorelačních vazeb o další veličiny, např. různé ukazatele socioekonomické a/nebo fyzické geografie. Za vědecky velmi významné považuje autor hlubší studium translace geografických vazeb objektů do prostoru parametrů ANN.



## 9 ZÁVĚR

Hlavním cílem práce bylo sestavit ANN pro potřeby predikčního modelování vybraných epidemiologických charakteristik pandemie COVID-19 na území ČR. K jeho splnění bylo potřeba modelovaná data nejprve analyzovat, navrhnout konkrétní architekturu ANN a představit způsob validace výsledného modelu.

Na zvolená modelovaná data, tj. počtu aktuálně nakažených na území ČR, byly aplikovány vybrané metody analýzy časových řad a datová transformace s cílem tvorby co nejrobustnějšího modelu schopného efektivního učení. Jednalo se o testování stacionarity ADF testem, využití autokorelace časových řad pomocí ACF ke stanovení velikosti posuvného okna a transformace dat Yeo-Johnsonovou mocninnou transformací.

Ze sekvenční povahy modelovaných dat a cíle predikčního modelování byla zvolena ANN využívající rekurentní GRU vrstvy a výstupní bayesovskou lineární vrstvu, díky čemuž jsou výsledné predikce doplněny mírou nejistoty. Rozebrané relevantní teoretické principy ANN jsou následně přeneseny do praktické implementace autorem vytvořené třídy B-GRU-PTSAR, která kromě samotné tvorby ANN spravuje také potřebné analýzy a metody validace výsledného modelu. Umožňuje také model sestavit zcela automatizovaně pomocí algoritmu CMA-ES, což umožňuje efektivnější tvorbu modelu a snižuje množství subjektivních rozhodnutí. Třída je sepsána obecně a umožňuje tvorbu podobných modelů i pro jiné účely.

Validace výsledného modelu proběhla vyhodnocením procesu automatické optimalizace hyperparametrů, analýzou procesu učení sítě napříč epochami učení a rozbořem přesnosti predikcí. Ty jsou porovnány vůči skutečnému stavu užitím grafu a několika mapových vizualizací. Schopnost modelu zachytit prostorové vazby byla kromě vizuální interpretace map testována metodou prostorové statistiky LISA analyzující prostorové shluky.

Cíleného výsledku práce bylo dosaženo výše zmíněnými metodami, přičemž výsledný model dosáhl uspokojivé přesnosti.

V kontextu celosvětového výzkumu lze za hlavní přednosti a přínos práce považovat využití bayesovských vrstev umožňující práci s mírou nejistoty predikcí a využití automatické optimalizace hyperparametrů při finalizaci modelu. Významné je také představení populárního přístupu k predikčnímu modelování (nejen) epidemiologických dat, který není v českém prostředí příliš používán. Z praktického hlediska je přínosná i třída B-GRU-PTSAR, jelikož poskytuje poměrně přímočarou cestu k tvorbě obdobných modelů i pro jiné úlohy. Z výsledků mohou potenciálně těžit relevantní zdravotnické, výzkumné a administrativní instituce.

## POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

ABBASIMEHR, Hossein a Reza PAKI. Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos, Solitons and Fractals* [online], 2020. 142(110511), 14 s. [cit. 2022-12-18]. ISSN 1873-2887. Dostupné z: doi:https://doi.org/10.1016/j.chaos.2020.110511

AGGARWAL, Charu C. *Artificial Intelligence: A Textbook*. 1st Edition. Cham: Springer Nature Switzerland, 2021. 503 s. ISBN 978-3-030-72356-9.

AGGARWAL, Charu C. *Neural Networks and Deep Learning: A Textbook*. 1st Edition. Cham: Springer Nature Switzerland, 2018. 520 s. ISBN 978-3-319-94462-3.

ALASSAFI, Madini O., Mutasem JARRAH a Reem ALOTAIBI. Time series predicting of COVID-19 based on deep learning. *Neurocomputing* [online], 2022. 468:335-344 [cit. 2022-10-28]. ISSN 1872-8286. Dostupné z: doi:https://doi.org/10.1016/j.neucom.2021.10.035

ALI, Peshawa J. M. a Rezhna H. FARAJ. Data Normalization and Standardization: A Technical Report. *Machine Learning Technical Reports* [online], 2014. 1(1):1-6 [cit. 2023-02-28]. Dostupné z: doi:https://doi.org/10.13140/RG.2.2.28948.04489

ARORA, Parul, Himanshu KUMAR a Bijaya K. PANIGRAHI. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, Solitons and Fractals* [online], 2020. 139(110017), 9 s. [cit. 2022-10-28]. ISSN 1873-2887. Dostupné z: doi:https://doi.org/10.1016/j.chaos.2020.110017

ARUNKUMAR, K. E., Dinesh V. KALAGA, Ch. M. S. KUMAR, Masahiro KAWAJI a Timothy M. BRENTA. Forecasting of COVID-19 using deep layer Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells. *Chaos, Solitons and Fractals* [online], 2021. 146(110861), 12 s. [cit. 2022-10-28]. ISSN 1873-2887. Dostupné z: doi:https://doi.org/10.1016/j.chaos.2021.110861

BI, Luning, Mohammad FILI a Guiping HU. COVID-19 forecasting and intervention planning using gated recurrent unit and evolutionary algorithm. *Neural Computing and Applications* [online], 2022. 34:17561-17579 [cit. 2022-10-29]. ISSN 1433-3058. Dostupné z: doi:https://doi.org/10.1007/s00521-022-07394-z

BIANCHI, Filippo M., Enrico MAIORINO, Michael C. KAMPPFMEYER, Antonello RIZZI a Robert JENSSEN. *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*. 1st Edition. Cham: Springer International Publishing, 2017. 81 s. ISBN 978-3-319-70337-4.

BISHOP, Christopher M. *Pattern Recognition and Machine Learning*. 1st Edition. New York: Springer-Verlag, 2006. 738 s. ISBN 978-0-387-31073-2.

BLUNDELL, Charles, Julien CORNEBISE, Koray KAVUKCUOGLU a Daan WIERSTRA. Weight Uncertainty in Neural Networks. *Proceedings of the 32nd International Conference on Machine Learning (ICML 2015)* [online], 2015. 10 s. [cit. 2022-12-01]. Dostupné z: doi:https://doi.org/10.48550/arXiv.1505.05424

BOJER, Casper S. Understanding machine learning-based forecasting methods: A decomposition framework and research opportunities. *International Journal of Forecasting* [online], 2022. 34(4):1555-1561 [cit. 2022-10-23]. ISSN 0169-2070. Dostupné z: doi:https://doi.org/10.1016/j.ijforecast.2021.11.003

BOTCHKAREV, Alexei. Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology. *Interdisciplinary Journal of Information, Knowledge, and Management* [online], 2019. 14:45-79 [cit. 2022-12-04]. ISSN 1555-1237. Dostupné z: doi:https://doi.org/10.28945/4184

CABRAS, Stefano. A Bayesian-Deep Learning Model for Estimating COVID-19 Evolution in Spain. *Mathematics* [online], 2021. 9(22):2921, 18 s. [cit. 2022-10-29]. ISSN 2227-7390. Dostupné z: doi:https://doi.org/10.3390/math9222921

CHANDRA, Rohitash a Yixuan HE. Bayesian neural networks for stock price forecasting before and during COVID-19 pandemic. *PLoS ONE* [online], 2021. 16(7):e0253217, 32 s. [cit. 2022-10-30]. ISSN 1932-6203. Dostupné z: doi:<https://doi.org/10.1371/journal.pone.0253217>

CHANDRA, Rohitash, Ayush JAIN a Divyanshu S. CHAUHAN. Deep learning via LSTM models for COVID-19 infection forecasting in India. *PLoS ONE* [online], 2022. 17(1):e0262708, 28 s. [cit. 2022-10-30]. ISSN 1932-6203. Dostupné z: doi:<https://doi.org/10.1371/journal.pone.0262708>

CHO, Kyunghyun, Bart van MERRIËNBOER, Caglar GULCEHRE, Dzmitry BAHDANAU, Fethi BOUGARES, Holger SCHWENK a Yoshua BENGIO. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of conference EMNLP 2014* [online], 2014. 15 s. [cit. 2022-11-27]. Dostupné z: doi:<https://doi.org/10.48550/arXiv.1406.1078>

DATHAN, Brahma a Sarnath RAMNATH. *Object-Oriented Analysis, Design and Implementation: An Integrated Approach*. 2nd Edition. New York: Springer International Publishing, 2015. 471 s. ISBN 978-3-319-24278-1.

DICKEY, David A. a Wayne A. FULLER. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association* [online], 1979. 74(366):427–431 [cit. 2023-02-10]. ISSN 1537-274X. Dostupné z: doi:<https://www.doi.org/10.2307/2286348> a url:[https://www.researchgate.net/publication/243644934\\_Distribution\\_of\\_the\\_Estimators\\_for\\_Autoregressive\\_Time\\_Series\\_With\\_a\\_Unit\\_Root](https://www.researchgate.net/publication/243644934_Distribution_of_the_Estimators_for_Autoregressive_Time_Series_With_a_Unit_Root)

FEURER, Matthias a Frank HUTTER. Hyperparameter Optimization. In: HUTTER, Frank, Lars KOTTHOFF a Joaquin VANSCHOREN, eds. *Automated Machine Learning: Methods, Systems, Challenges*. 1st Edition. Cham: Springer Nature Switzerland, 2019. s. 3–33. ISBN 978-3-030-05317-8.

FLORES, João H. F., Paulo M. ENGEL a Rafael C. PINTO. Autocorrelation and partial autocorrelation functions to improve neural networks models on univariate time series forecasting. *International Joint Conference on Neural Networks* [online], 2012. 8 s. [cit. 2023-02-13]. ISSN 2161-4407. Dostupné z: doi:<https://doi.org/10.1109/IJCNN.2012.6252470> url:[https://www.researchgate.net/publication/261464969\\_Autocorrelation\\_and\\_partial\\_autocorrelation\\_functions\\_to\\_improve\\_neural\\_networks\\_models\\_on\\_univariate\\_time\\_series\\_forecasting](https://www.researchgate.net/publication/261464969_Autocorrelation_and_partial_autocorrelation_functions_to_improve_neural_networks_models_on_univariate_time_series_forecasting)

FRITZ, Cornelius, Emilio DORIGATTI a David RÜGAMER. Combining graph neural networks and spatio-temporal disease models to improve the prediction of weekly COVID-19 cases in Germany. *Scientific reports* [online], 2022, 12:3930, 18 s. [cit. 2022-10-29]. ISSN 2045-2322. Dostupné z: doi:<https://doi.org/10.1038/s41598-022-07757-5>

GOAN, Ethan a Clinton FOOKES. Bayesian Neural Networks: An Introduction and Survey. In: MENGERSEN, Kerrie L., Pierre PUDLO a Christian P. ROBERT, eds. *Case Studies in Applied Bayesian Data Science: CIRM Jean-Morlet Chair, Fall 2018*. 1st Edition. Cham: Springer Nature Switzerland, 2020. s. 45–87. ISBN 978-3-030-42552-4.

GOODFELLOW, Ian, Yoshua BENGIO a Aaron COURVILLE. *Deep Learning*. 1st Edition. Cambridge, MA: MIT Press, 2016. 800 s. ISBN 978-0-262-03561-3.

HANSEN, Nikolaus. *The CMA Evolution Strategy: A Tutorial* [online]. 2016. 37 s. [cit. 2023-02-10]. Dostupné z: doi:<https://doi.org/10.48550/arXiv.1604.00772>

HAYKIN, Simon. *Neural Networks and Learning Machines*. 3rd Edition. New Jersey: Pearson Educational, 2009. 936 s. ISBN 978-0-13-147139-9.

HAZRA, Avijit. Using the Confidence Interval Confidently. *Journal of Thoracic Disease* [online], 2017. 9(10):4125–4130 [cit. 2023-03-25]. ISSN 2077-6624. Dostupné z: doi:<http://dx.doi.org/10.21037/jtd.2017.09.14>

HERRANZ, Edward. Unit Root Tests. *Wiley Interdisciplinary Reviews: Computational Statistics* [online], 2017. 9(3):e1396, 20 s. [cit. 2023-02-12]. ISSN 1939-0068. Dostupné z: doi:<https://doi.org/10.1002/wics.1396> a url:[https://www.researchgate.net/publication/315816808\\_Unit\\_root\\_tests](https://www.researchgate.net/publication/315816808_Unit_root_tests)

HOCHREITER, Sepp a Jürgen SCHMIDHUBER. Long Short-Term Memory. *Neural Computation* [online], 1997. 9(8):1735–1780 [cit. 2022-11-27]. ISSN 0899-7667. Dostupné z: doi:https://doi.org/10.1162/neco.1997.9.8.1735

HOMAYOUNI, Hajar, Sudipto GHOSH, Indrakshi RAY, Shlok GONDALIA, Jerry DUGGAN a Michael G. KAHN. An Autocorrelation-based LSTM-Autoencoder for Anomaly Detection on Time-Series Data. *IEEE International Conference on Big Data* [online], 2020. s. 5068–5077 [cit. 2023-02-13]. ISSN 2573-2978. Dostupné z: doi:https://doi.org/10.1109/BigData50022.2020.9378192 url:https://www.researchgate.net/publication/350197384\_An\_Autocorrelation-based\_LSTM-Autoencoder\_for\_Anomaly\_Detection\_on\_Time-Series\_Data

HUTTER, Frank, Lars KOTTHOFF a Joaquin VANSCHOREN. Preface. In: HUTTER, Frank, Lars KOTTHOFF a Joaquin VANSCHOREN, eds. *Automated Machine Learning: Methods, Systems, Challenges*. 1st Edition. Cham: Springer Nature Switzerland, 2019. s. ix–xi. ISBN 978-3-030-05317-8.

HYNDMAN, Robin J. a George ATHANASOPOULOS. *Forecasting: Principles and Practice*. 3rd Edition. Melbourne: OTexts, 2021. 442 s. ISBN 978-0987507136.

JANOUT, Vladimír a Jana JANOUTOVÁ. *Medicína založená na důkazu a klinická epidemiologie*. 1. edice. Praha: Grada Publishing, 2021. 200 s. ISBN 978-80-271-3076-4.

JARKOVSKÝ, Jiří, Klára BENEŠOVÁ, Ondřej MÁJEK a Ladislav DUŠEK. Dokumentace k modelu ÚZIS ČR pro predikce počtu hospitalizovaných pacientů a pacientů vyžadujících intenzivní péči. In: *Předávací úložiště ÚZIS ČR: Dokumentace prediktivních modelů* [online]. Praha: ÚZIS ČR, 2021. 14 s. [cit. 2022-11-02]. Dostupné z: https://share.uzis.cz/s/cmFHjc4jbqPBAER

JINDAL, Ishan, Matthew NOKLEBY, Daniel PRESSEL, Xuewen CHEN a Harpreet SINGH. Deep Neural Networks for Corrupted Labels. In: PEDRYCZ, Witold a Shyi-Ming CHEN, eds. *Deep Learning: Concepts and Architectures*. 1st Edition. Cham: Springer Nature Switzerland, 2020. s. 211–235. ISBN 978-3-030-31755-3.

JOSHI, Ameet V. *Machine Learning and Artificial Intelligence*. 2nd Edition. Cham: Springer Nature Switzerland, 2023. 284 s. ISBN 978-3-031-12281-1.

JOSPIN, Laurent V., Hamid LAGA, Farid BOUSSAID, Wray BUNTINE a Mohammed BENNAMOUN. Hands-on Bayesian Neural Networks: A Tutorial for Deep Learning Users. *IEEE Computational Intelligence Magazine* [online], 2022. 17(2):29–48 [cit. 2022-11-29]. ISSN 1556-6048. Dostupné z: doi:https://doi.org/10.1109/MCI.2022.3155327

KAČÍRKOVÁ, Tereza. *Důsledky pandemie COVID-19 na mezinárodní mobilitu v Evropě* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2022. 81 s. [cit. 2022-11-04]. Magisterská práce. Vedoucí práce Mgr. Vít Pászto, Ph.D. Dostupné z: url:https://www.geoinformatics.upol.cz/dprace/magisterske/kacirkova22/downloads/kacirkova\_2\_2\_dp.pdf

KAMALOV, Firuz, Khairan RAJAB, Aswani K. CHERUKURI, Ashraf ELNAGAR a Murodbek SAFARALIEV. Deep learning for Covid-19 forecasting: State-of-the-art review. *Neurocomputing* [online], 2022. 511:142–154 [cit. 2022-10-26]. ISSN 1872-8286. Dostupné z: doi:https://doi.org/10.1016/j.neucom.2022.09.005

KAUL, Piyush a Brejesh LALL. Theoretical Characterization of Deep Neural Networks. In: PEDRYCZ, Witold a Shyi-Ming CHEN, eds. *Deep Learning: Concepts and Architectures*. 1st Edition. Cham: Springer Nature Switzerland, 2020. s. 25–63. ISBN 978-3-030-31755-3.

KHAN, Riaz U., Sultan ALMAKDI, Mohammed ALSHEHRI, Rajesh KUMAR, Ikram ALI, Sardar M. HUSSAIN, Amin UI HAQ, Inayat KHAN, Aman ULLAH a Muhammad I. UDDIN. Probabilistic Approach to COVID-19 Data Analysis and Forecasting Future Outbreaks Using a Multi-Layer Perceptron Neural Network. *Diagnostics* [online], 2022. 12(10):2539, 25 s. [cit. 2022-10-30]. ISSN 2075-4418. Dostupné z: doi:https://doi.org/10.3390/diagnostics12102539

KINGMA, Diederik P. a Jimmy BA. Adam: A Method for Stochastic Optimization. *3rd International Conference for Learning Representations* [online], 2015. 15 s. [cit. 2022-12-29]. Dostupné z: doi:https://doi.org/10.48550/arXiv.1412.6980

KITTNAR, Otomar, Kateřina JANDOVÁ, Eduard KURIŠČÁK, Miloš LANGMEIER, Dana MAREŠOVÁ, Mikuláš MLČEK, Jaromír MYSLIVEČEK, Jaroslav POKORNÝ, Vladimír RILJAK a Stanislav TROJAN. *Lékařská fyziologie*. 2. vydání. Praha: Grada Publishing, 2020. 752 s. ISBN 978-80-247-1963-4.

KRUSE, Rudolf, Sanaz MOSTAGHIM, Christian BORGELT, Christian BRAUNE a Matthias STEINBRECHER. *Computational Intelligence: A Methodological Introduction*. 3rd Edition. Cham: Springer Nature Switzerland, 2022. 653 s. ISBN 978-3-030-42226-4.

LUCAS, Benjamin, Behzad VAHEDI a Morteza KARIMZADEH. A spatiotemporal machine learning approach to forecasting COVID-19 incidence at the county level in the USA. *International Journal of Data Science and Analytics* [online], 2022. 20 s. [cit. 2022-10-29]. ISSN 2364-4168. Dostupné z: doi:https://doi.org/10.1007/s41060-021-00295-9

LUZ, Leonard. *Designing a Thematic Atlas on the Geospatial Impact of COVID-19* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2021. 76 s. [cit. 2022-11-04]. Magisterská práce. Vedoucí práce RNDr. Alena Vondráková, Ph.D., LL.M. Dostupné z: url:https://www.geoinformatics.upol.cz/dprace/magisterske/luz21/documents/thesis\_text\_lluz.pdf

MÁJEK, Ondřej, Ondřej NGO, Renata CHLOUPKOVÁ, Jiří JARKOVSKÝ, Monika AMBROŽOVÁ, Barbora BUDÍKOVÁ, Jan KOUŘIL, Tomáš PAVLÍK a Ladislav DUŠEK. Dokumentace k epidemiologickému modelu ÚZIS ČR pro krátkodobé simulace. In: *Předávací úložiště ÚZIS ČR: Dokumentace prediktivních modelů* [online]. Praha: ÚZIS ČR, 2021. 20 s. [cit. 2022-10-30]. Dostupné z: url:https://share.uzis.cz/s/cmFHjc4jbqPBAER

MÁJEK, Ondřej, Ondřej NGO, Renata CHLOUPKOVÁ, Jiří JARKOVSKÝ, Tomáš PAVLÍK, Martin KOMENDA a Ladislav DUŠEK. Dokumentace k epidemiologickému modelu ÚZIS ČR pro dlouhodobé simulace. In: *Předávací úložiště ÚZIS ČR: Dokumentace prediktivních modelů* [online]. Praha: ÚZIS ČR, 2021. 9 s. [cit. 2022-11-02]. Dostupné z: url:https://share.uzis.cz/s/cmFHjc4jbqPBAER

MATOŠKA, Michal. *Návrh a analýza epidemiologického modelu pro predikci vývoje COVID-19* [online]. Ostrava: Fakulta elektrotechniky a informatiky VŠB Technické univerzity Ostrava, 2022. 72 s. [cit. 2022-11-04]. Bakalářská práce. Vedoucí práce Ing. Jan Kubíček, Ph.D. Dostupné z: url:https://dspace.vsb.cz/bitstream/handle/10084/147368/MAT0397\_FEI\_B0714A060016\_2022.pdf?sequence=1&isAllowed=y

MCCULLOCH, Warren S. a Walter PITTS. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 1943. 5:115–133. ISSN 0007-4985. Dostupné z: doi:https://www.doi.org/10.1007/bf02478259

MILLS, Kerry R., ed. *Oxford Textbook of Clinical Neurophysiology*. 1st Edition. Oxford: Oxford University Press, 2017. 480 s. ISBN 978-0-19-968839-5.

MOEIN, Shiva, Niloofar NICKAEEN, Amir ROOINTAN, Niloofar BORHANI, Zarifeh HEIDARY, Shaghayegh H. JAVANMARD, Jafar GHASARI a Yousof GHEISARI. Inefficiency of SIR models in forecasting COVID-19 epidemic: a case study of Isfahan. *Scientific reports* [online], 2021. 41(4725), 9 s. [cit. 2022-10-23]. ISSN 2045-2322. Dostupné z: doi:https://doi.org/10.1038/s41598-021-84055-6

MOHIMONT, Lucas, Amine CHEMCHAM, François ALIN, Michael KRAJECKI a Luiz A. STEFFENEL. Convolutional neural networks and temporal CNNs for COVID-19 forecasting in France. *Applied Intelligence* [online], 2021. 51:8784–8809 [cit. 2022-10-26]. ISSN 1573-7497. Dostupné z: doi:https://doi.org/10.1007/s10489-021-02359-6

MONTGOMERY, Douglas C., Cheryl L. JENNINGS a Murat KULAHCI. *Introduction to Time Series Analysis and Forecasting*. 2nd Edition. Hoboken: Wiley Interscience, 2015. 672 s. ISBN 978-1118745113.

MUÑOZ-ORGANERO, Mario a Paula QUEIPO-ÁLVAREZ. Deep Spatiotemporal Model for COVID-19 Forecasting. *Sensors* [online], 2022. 22(9):3519, 17 s. [cit. 2022-10-26]. ISSN 1424-8220. Dostupné z: doi:https://doi.org/10.3390/s22093519

NEVTÍPILOVÁ, Veronika. *Testování neuronových sítí pro prostorovou interpolaci v softwaru GRASS GIS* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2013. 47 s. [cit. 2022-11-04]. Bakalářská práce. Vedoucí práce Mgr. Justyna Pastwa. Dostupné z: [url:https://www.geoinformatics.upol.cz/dprace/bakalarske/nevtipilova13/text\\_prace.pdf](http://www.geoinformatics.upol.cz/dprace/bakalarske/nevtipilova13/text_prace.pdf)

NGUYEN, Giang, Stefan DLUGOLINSKY, Martin BOBÁK, Viet TRAN, Álvaro LÓPEZ-GARCÍA, Ignacio HEREDIA, Peter MALÍK a Ladislav HLUCHÝ. Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review* [online], 2019, 52:77–124 [cit. 2022-11-11]. ISSN 1573-7462. Dostupné z: [doi:https://doi.org/10.1007/s10462-018-09679-z](https://doi.org/10.1007/s10462-018-09679-z)

NIRAULA, Poshan, Jorge MATEU a Somnath CHAUDHURI. A Bayesian machine learning approach for spatio-temporal prediction of COVID-19 cases. *Stochastic Environmental Research and Risk Assessment* [online], 2022. 36:2265–2283 [cit. 2022-10-30]. ISSN 1436-3259. Dostupné z: [doi:https://doi.org/10.1007/s00477-021-02168-w](https://doi.org/10.1007/s00477-021-02168-w)

NIXON, Kristen, Sonia JINDAL, Felix PARKER, Maximilian MARSHALL, Nicholas G. REICH, Kimia GHOBADI, Elizabeth C. LEE, Shaun TRUELOVE a Lauren GARDNER. Real-time COVID-19 forecasting: challenges and opportunities of model performance and translation. *The Lancet: Digital Health* [online], 2022. 4(10):699–701. [cit. 2022-10-30]. ISSN 2589-7500. Dostupné z: [doi:https://doi.org/10.1016/S2589-7500\(22\)00167-4](https://doi.org/10.1016/S2589-7500(22)00167-4)

OSADEBEY, Michael, Marius PEDERSEN, Douglas ARNOLD a Katrina WENDEL-MITORAJ. Local Indicators of Spatial Autocorrelation (LISA): Application to Blind Noise-Based Perceptual Quality Metric Index for Magnetic Resonance Images. *Journal of Imaging* [online], 2019. 5(20), 23 s. [cit. 2023-03-27]. ISSN 2313-433X. Dostupné z: [doi:http://dx.doi.org/10.3390/jimaging5010020](http://dx.doi.org/10.3390/jimaging5010020)

PATARWAL, Preeti, Ram P. SINGH a Rajni BALA. Financial and Non-Stationary Time Series Forecasting using LSTM Recurrent Neural Network for Short and Long Horizon. *International Conference on Computing, Communication, and Networking Technologies* [online]. 2019. 10, 7 s. [cit. 2023-02-12]. ISSN 2473-7674. Dostupné z: [doi:https://doi.org/10.1109/ICCCNT45670.2019.8944624](https://doi.org/10.1109/ICCCNT45670.2019.8944624)

PORTI SUÁREZ, Anna. *User Evaluation of COVID-19 Dashboards* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2022. 71 s. [cit. 2022-11-04]. Magisterská práce. Vedoucí práce RNDr. Stanislav Popelka, Ph.D. Dostupné z: [url:https://www.geoinformatics.upol.cz/dprace/magisterske/PortiSuarez22/documents/Text.pdf](https://www.geoinformatics.upol.cz/dprace/magisterske/PortiSuarez22/documents/Text.pdf)

PREM, Kiesha, Yang LIU, Timothy W. RUSSELL, Adam J. KUCHARSKI, Rosalind M. EGGO, Nicholas DAVIES, Mark JIT a Petra KLEPAC. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet: Public Health* [online], 2020. 5(5):261–270 [cit. 2022-11-02]. ISSN 2468-2667. Dostupné z: [doi:https://doi.org/10.1016/S2468-2667\(20\)30073-6](https://doi.org/10.1016/S2468-2667(20)30073-6)

PREM, Kiesha, Alex R. COOK a Mark JIT. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology* [online], 2017. 13(9):e1005697, 21 s. [cit. 2022-11-02]. ISSN 1553-7358. Dostupné z: [doi:https://doi.org/10.1371/journal.pcbi.1005697](https://doi.org/10.1371/journal.pcbi.1005697)

PŘIBYLOVÁ, Lenka a Veronika HAJNOVÁ. Model epidemie COVID-19. In: *Předávací úložiště ÚZIS ČR* [online]. Brno: Ústav matematiky a statistiky Přírodovědecké fakulty Masarykovy Univerzity, 2021. 16 s. [cit. 2022-11-03]. Dostupné z: <https://share.uzis.cz/s/cmFHjc4jbqPBAER>

QU, Rong. Recent Developments of Automated Machine Learning and Search Techniques. In: PILLAY, Nelishia a Rong QU, eds. *Automated Design of Machine Learning and Search Algorithms*. 1st Edition. Cham: Springer Nature Switzerland, 2021. s. 1–9. ISBN 978-3-030-72068-1.

RGUIBI, Mohamed A., Najem MOUSSA, Abdellah MADANI, Abdessadak AAROUD a Khalid ZINE-DINE. Forecasting Covid-19 Transmission with ARIMA and LSTM Techniques in Morocco. *Springer Nature: Computer Science* [online], 2022. 3:133, 14 s. [cit. 2022-10-26]. ISSN 2661-8907. Dostupné z: [doi:https://doi.org/10.1007/s42979-022-01019-x](https://doi.org/10.1007/s42979-022-01019-x)

RODRIGUES, Helena S. Application of SIR epidemiological model: new trends. *International Journal of Applied Mathematics and Informatics* [online], 2016. 10:92–97 [cit. 2022-10-23]. ISSN 2074-1278. Dostupné z: doi:<https://doi.org/10.48550/arXiv.1611.02565>

ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 1958. 65(6):386–408. ISSN 0033-295X. Dostupné z: doi:<https://doi.org/10.1037/h0042519>

SADÍLEK, Martin. *Vymezení výtvarného stylu tematických map metodami machine learning* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2021. 92 s. [cit. 2022-11-04]. Magisterská práce. Vedoucí práce prof. RNDr. Vít Voženilek, CSc. Dostupné z: url:<https://www.geoinformatics.upol.cz/dprace/magisterske/sadilek21/text.pdf>

SAHAI, Alok K., Namita RATH, Vishal SOOD a Manvendra P. SINGH. ARIMA modelling & forecasting of COVID-19 in top five affected countries. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* [online], 2020. 14(5):1419–1427 [cit. 2022-10-24]. ISSN 1871-4021. Dostupné z: doi:<https://doi.org/10.1016/j.dsx.2020.07.042>

SAID, Said E. a David A. DICKEY. Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order. *Biometrika* [online], 1984. 71(3):599–607 [cit. 2023-02-12]. ISSN 1464-3510. Dostupné z: doi:<https://doi.org/10.2307/2336570> a url:<https://www.ssc.wisc.edu/~bhansen/718/SaidDickey1984.pdf>

SALIAJ, Lorena a Eugenia NISSI. An Application of Neural Networks to Predict COVID-19 Cases in Italy. *Engineering proceedings* [online], 2022. 18(1):11, 8 s. [cit. 2022-10-29]. ISSN 2673-4591. Dostupné z: doi:<https://doi.org/10.3390/engproc2022018011>

SEIDL, Zdeněk. *Neurologie pro studium i praxi*. 2. vydání. Praha: Grada Publishing, 2015. 384 s. ISBN 978-80-247-5247-1.

SILK, Daniel S., Veronica E. BOWMAN, Daria SEMOCHKINA, Ursula DALRYMPLE a Dave C. WOODS. Uncertainty quantification for epidemiological forecasts of COVID-19 through combinations of model predictions. *Statistical Methods in Medical Research* [online], 2022. 31(9):1778–1789 [cit. 2022-10-30]. ISSN 1477-0334. Dostupné z: doi:<https://doi.org/10.1177/09622802221109523> a url:[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9272045/pdf/10.1177\\_09622802221109523.pdf](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9272045/pdf/10.1177_09622802221109523.pdf)

STEINBACH, Peter, Felicita GERNHARDT, Mahnoor TANVEER, Steve SCHMERLER a Sebastian STARKE. Machine Learning State-of-the-Art with Uncertainties. *ML Evaluation Standards Workshop at ICLR 2022* [online], 2022. 9 s. [cit. 2022-10-23]. Dostupné z: doi:<https://doi.org/10.48550/arXiv.2204.05173>

ŠALENÁ, Hana. *Predikce účinnosti opatření proti šíření COVID-19 v ČR* [online]. Brno: Fakulta sociálních studií Masarykovy Univerzity, 2021. 149 s. [cit. 2022-11-04]. Magisterská práce. Vedoucí práce doc. Ing. Josef Procházka, Ph.D. Dostupné z: url:[https://is.muni.cz/th/q2i1q/DP\\_Salena\\_Hana\\_finalni\\_verze.pdf](https://is.muni.cz/th/q2i1q/DP_Salena_Hana_finalni_verze.pdf)

ŠIROKÝ, Jan. *Tvořme a publikujeme odborné texty: Nejen pro ekonomy a manažery*. 1. edice, Brno: Computer Press, 2011. 208 s. ISBN 978-80-251-3510-5.

TABATABA, Farzaneh S., Prithwish CHAKRABORTY, Naren RAMAKRISHNAN, Srinivasan VENKATRAMANAN, Jiangzhuo CHEN, Bryan LEWIS a Madhav MARATHE. A framework for evaluating epidemic forecasts. *BMC Infectious Diseases* [online], 2017. 17(345), 27 s. [cit. 2023-03-29]. ISSN 1471-2334. Dostupné z: doi:<https://doi.org/10.1186/s12879-017-2365-1>

TAN, Cia Vei, Sarbhan SINGH, Chee Heng LAI, Ahmed S. SYAFIQ MD ZAMRI, Sarat Chandra DASS, Tahir BIN ARIS, Hishamshah M. IBRAHIM a Balvinder S. GILL. Forecasting COVID-19 Case Trends Using SARIMA Models during the Third Wave of COVID-19 in Malaysia. *International Journal of Environmental Research and Public Health* [online], 2022. 19(3):1504–1516 [cit. 2022-10-24]. ISSN 1660-4601. Dostupné z: doi:<https://doi.org/10.3390/ijerph19031504>

TYL, Erik. *Neuronové sítě při analýze výsledků voleb* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2013. 74 s. [cit. 2022-11-04]. Magisterská práce. Vedoucí práce doc. Mgr. Jirí Dvorský, Ph.D. Dostupné z: [url:https://docplayer.cz/44646401-Neuronove-site-pri-analyze-vysledku-voleb.html](https://docplayer.cz/44646401-Neuronove-site-pri-analyze-vysledku-voleb.html)

URL1: PREM, Kiesha. *kieshaprem/covid19-agestructureSEIR-wuhan-social-distancing* [online]. GitHub, 2020. [cit. 2023-04-04]. Dostupné z: <https://github.com/kieshaprem/covid19-agestructureSEIR-wuhan-social-distancing>

URL2: COVID-19 v ČR: *Otevřené datové sady a sady ke stažení* [online]. Onemocnění aktuálně od MZČR. [cit. 2023-04-04]. Dostupné z: <https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19>

URL3: ArcČR® 4.1: *Vybraná administrativní a statistická data o České republice* [online]. ARCDATA PRAHA. [cit. 2023-04-04]. Dostupné z: <https://www.arcdata.cz/produkty/geograficka-data/arccr-4>

URL4: *Stationarity and detrending (ADF/KPSS)* [online]. Statsmodels. [cit. 2023-04-04]. Dostupné z: [https://www.statsmodels.org/dev/examples/notebooks/generated/stationarity\\_detrending\\_adf\\_kpss.html](https://www.statsmodels.org/dev/examples/notebooks/generated/stationarity_detrending_adf_kpss.html)

URL5: *sklearn.preprocessing.StandardScaler* [online]. Scikit-learn: Machine Learning in Python. [cit. 2023-04-04]. Dostupné z: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>

URL6: *6.3. Preprocessing data: 6.3.1.3 Scaling data with outliers* [online]. Scikit-learn: Machine Learning in Python. [cit. 2023-04-04]. Dostupné z: <https://scikit-learn.org/stable/modules/preprocessing.html#scaling-data-with-outliers>

URL7: *sklearn.preprocessing.RobustScaler* [online]. Scikit-learn: Machine Learning in Python. [cit. 2023-04-04]. Dostupné z: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html>

URL8: *sklearn.preprocessing.PowerTransformer* [online]. Scikit-learn: Machine Learning in Python. [cit. 2023-04-04]. Dostupné z: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.PowerTransformer.html>

URL9: *6.3 Preprocessing data: 6.3.2.2 Mapping to a Gaussian distribution* [online]. Scikit-learn: Machine Learning in Python. [cit. 2023-04-04]. Dostupné z: <https://scikit-learn.org/stable/modules/preprocessing.html#mapping-to-a-gaussian-distribution>

URL10: *Compare the effect of different scalers on data with outliers* [online]. Scikit-learn: Machine Learning in Python. [cit. 2023-04-04]. Dostupné z: [https://scikit-learn.org/stable/auto\\_examples/preprocessing/plot\\_all\\_scaling.html](https://scikit-learn.org/stable/auto_examples/preprocessing/plot_all_scaling.html)

URL11: *PyTorch Docs: Linear* [online]. PyTorch, 2023. [cit. 2023-04-04]. Dostupné z: <https://pytorch.org/docs/stable/generated/torch.nn.Linear.html>

URL12: *PyTorch Docs: torch.Tensor* [online]. PyTorch, 2023. [cit. 2023-04-04]. Dostupné z: <https://pytorch.org/docs/stable/tensors.html>

URL13: vcubingx. *Why are neural networks so effective?* [online]. YouTube, 2021. [cit. 2023-04-04]. Dostupné z: [https://youtu.be/-at7SLoVK\\_I?t=395](https://youtu.be/-at7SLoVK_I?t=395)

URL14: *PyTorch Docs: RNN* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: <https://pytorch.org/docs/stable/generated/torch.nn.RNN.html>

URL15: *PyTorch Docs: GRU* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: <https://pytorch.org/docs/stable/generated/torch.nn.GRU.html>

URL16: ESPOSITO, Piero. *Weight a priori and a posteriori sampler: class PriorWeightDistribution* [online]. GitHub, 2020. [cit. 2023-04-05]. Dostupné z: <https://github.com/piEsposito/blitz-bayesian-deep-learning/blob/master/doc/samplers.md#class-PriorWeightDistribution>



- URL17: *PyTorch Docs: MSELoss* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: <https://pytorch.org/docs/stable/generated/torch.nn.MSELoss.html>
- URL18: *PyTorch Docs: L1Loss* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: <https://pytorch.org/docs/stable/generated/torch.nn.L1Loss.html>
- URL19: ESPOSITO, Piero. *Blitz - Bayesian Layers in Torch Zoo: It is also true that there is complexity cost function differentiable along its variables* [online]. GitHub, 2022. [cit. 2023-04-05]. Dostupné z: <https://github.com/piEposito/blitz-bayesian-deep-learning#It-is-also-true-that-there-is-complexity-cost-function-differentiable-along-its-variables>
- URL20: ESPOSITO, Piero. *blitz-bayesian-deep-learning/blitz/utils/variational\_estimator.py* [online]. GitHub, 2021. [cit. 2023-04-05]. Dostupné z: [https://github.com/piEposito/blitz-bayesian-deep-learning/blob/master/blitz/utils/variational\\_estimator.py](https://github.com/piEposito/blitz-bayesian-deep-learning/blob/master/blitz/utils/variational_estimator.py)
- URL21: ESPOSITO, Piero. *blitz-bayesian-deep-learning/blitz/losses/kl\_divergence.py* [online]. GitHub, 2020. [cit. 2023-04-05]. Dostupné z: [https://github.com/piEposito/blitz-bayesian-deep-learning/blob/master/blitz/losses/kl\\_divergence.py](https://github.com/piEposito/blitz-bayesian-deep-learning/blob/master/blitz/losses/kl_divergence.py)
- URL22: *PyTorch Docs: torch.Optim* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>
- URL23: *Optuna Docs: optuna.samplers* [online]. Optuna, 2018. [cit. 2023-04-05]. Dostupné z: <https://optuna.readthedocs.io/en/stable/reference/samplers/index.html>
- URL24: *PyTorch Docs: torch.optim.lr\_scheduler* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: [https://pytorch.org/docs/stable/\\_modules/torch/optim/lr\\_scheduler.html#ExponentialLR](https://pytorch.org/docs/stable/_modules/torch/optim/lr_scheduler.html#ExponentialLR)
- URL25: *PyTorch Docs: torch.load* [online]. PyTorch, 2023. [cit. 2023-04-05]. Dostupné z: <https://pytorch.org/docs/stable/generated/torch.load.html>
- VEGA, Roberto, Leonardo FLORES a Russell GREINER. SIMLR: Machine Learning inside the SIR Model for COVID-19 Forecasting. *Forecasting* [online], 2022. 4(1):72–94 [cit. 2022-10-24]. ISSN 2571-9394. Dostupné z: doi:<https://doi.org/10.3390/forecast4010005>
- VERMA, Hanuman, Saurav MANDAL a Akshansh GUPTA. Temporal deep learning architecture for prediction of COVID-19 cases in India. *Expert Systems With Applications* [online], 2022. 195:116611, 12 s. [cit. 2022-10-28]. ISSN 1873-6793. Dostupné z: doi:<https://doi.org/10.1016/j.eswa.2022.116611>
- VIERING, Tom a Marco LOOG. The Shape of Learning Curves: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online], In print. 20 s. [cit. 2023-03-21]. ISSN 1939-3539. Dostupné z: doi:<https://doi.org/10.1109/TPAMI.2022.3220744>
- VIRILI, Francesco a Bernd FREISLEBEN. Nonstationarity and Data Preprocessing for Neural Network Predictions of an Economic Time Series. *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks* [online], 2000. 5:129–134 [cit. 2023-02-12]. ISSN 2161-4407. Dostupné z: doi:<https://doi.org/10.1109/IJCNN.2000.861446>
- VÍTKOVÁ, Tereza. *Dostupnost kapacit zdravotnické péče během pandemie COVID-19* [online]. Olomouc: Přírodovědecká fakulta Univerzity Palackého v Olomouci, 2022. 49 s. [cit. 2022-11-04]. Bakalářská práce. Vedoucí práce Mgr. Vít Pászto, Ph.D. Dostupné z: url:[https://www.geoinformatics.upol.cz/dprace/bakalarske/vitkova22/download/BP\\_Vitkova.pdf](https://www.geoinformatics.upol.cz/dprace/bakalarske/vitkova22/download/BP_Vitkova.pdf)
- VOŽENÍLEK, Vít, Jaromír KAŇOK, Jan D. BLÁHA, Zdena DOBEŠOVÁ, Tomáš HUDEČEK, Magdaléna KOZÁKOVÁ a Zuzana NĚMCOVÁ. *Metody tematické kartografie: Vizualizace prostorových jevů*. 1. edice. Olomouc: Univerzita Palackého v Olomouci, 2011. 216 s. ISBN 978-80-244-2790-4.
- WAN, Xiang, Wenqian WANG, Jiming LIU a Tiejun TONG. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology* [online], 2014. 14(135), 13 s. [cit. 2023-04-15]. ISSN 1471-2288. Dostupné z: doi:<https://doi.org/10.1186/1471-2288-14-135>

WANG, Qi, Yue MA, Kun ZHAO a Yingjie TIAN. A Comprehensive Survey of Loss Functions in Machine Learning. *Annals of Data Science* [online], 2022. 9:187–212 [cit. 2022-12-04]. ISSN 2198-5812. Dostupné z: doi:<https://doi.org/10.1007/s40745-020-00253-5> url: <https://faculty.ist.psu.edu/vhonavar/Courses/ds310/lossfunc.pdf>

WARING, Jonathan, Charlotta LINDVALL a Renato UMETON. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. *Artificial Intelligence in Medicine* [online], 2020. 104(101822), 12 s. [cit. 2023-02-09]. ISSN 1873-2860. Dostupné z: doi:<https://doi.org/10.1016/j.artmed.2020.101822>

WIBAWA, Aji P., Agung B. PUTRA UTAMA, Hakkun ELMUNSYAH, Utomo PUJIANTO, Felix A. DWIYANTO a Leonel HERNANDEZ. Time-series analysis with smoothed Convolutional Neural Network. *Journal of Big Data* [online], 2022. 9(44), 18 s. [cit. 2022-10-26]. ISSN 2196-1115. Dostupné z: doi:<https://doi.org/10.1186/s40537-022-00599-y>

WIEMKEN, Timothy L. a Robert R. KELLEY. Machine Learning in Epidemiology and Health Outcomes Research. *Annual Review of Public Health* [online], 2020. 41:21–36 [cit. 2022-10-23]. ISSN 1545-2093. Dostupné z: doi:<https://doi.org/10.1146/annurev-publhealth-040119-094437>

YEO, In-Kwon a Richard A. JOHNSON. A new family of power transformations to improve normality or symmetry. *Biometrika* [online], 2000. 87(4):954–959 [cit. 2023-02-28]. ISSN 1464-3510. Dostupné z: doi:<https://doi.org/10.1093/biomet/87.4.954>

ZAIN, Zuhaira M. a Nazik M. ALTURKI. COVID-19 Pandemic Forecasting Using CNN-LSTM: A Hybrid Approach. *Journal of Control Science and Engineering* [online], 2021. ID 8785636, 23 s. [cit. 2022-10-26]. ISSN 1687-5257. Dostupné z: doi:<https://doi.org/10.1155/2021/8785636>

ZAKI, Mohammed J. a Wagner MEIRA, Jr. *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*. 2nd Edition. Cambridge, UK: Cambridge University Press, 2020. 776 s. ISBN 978-1108473989.

ZENG, Jiaming, Adam LESNIKOWSKI a Jose M. ALVAREZ. The Relevance of Bayesian Layer Positioning to Model Uncertainty in Deep Bayesian Active Learning. *Third workshop on Bayesian Deep Learning (NeurIPS 2018)* [online], 2018. 6 [cit. 2022-11-22]. Dostupné z: doi:<https://doi.org/10.48550/arXiv.1811.12535>

ZHANG, Guoqiang, Butje E. PATUWO a Michael Y. HU. Forecasting with artificial neural networks: The State of Art. *International Journal of Forecasting* [online], 1998. 14(1):35–62 [cit. 2023-02-14]. ISSN 1872-8200. Dostupné z: doi:[https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7) a url:[https://www.researchgate.net/publication/222489987\\_Forecasting\\_With\\_Artificial\\_Neural\\_Networks\\_The\\_State\\_of\\_the\\_Art](https://www.researchgate.net/publication/222489987_Forecasting_With_Artificial_Neural_Networks_The_State_of_the_Art)

ZHOU, Zhi-Hua, Yang YU a Chao QIAN. *Evolutionary Learning: Advances in Theories and Algorithms*. 1st Edition. Singapur: Springer Nature Singapore, 2019. 376 s. ISBN 978-981-13-5955-2.

ZHOU, Kun, Wen Yong WANG, Teng HU a Chen Huang WU. Comparison of Time Series Forecasting Based on Statistical ARIMA Model and LSTM with Attention Mechanism. *Journal of Physics: Conference Series* [online], 2020. 1631(012141), 10 s. [cit. 2022-10-26]. ISSN 1742-6596. Dostupné z: doi:<https://doi.org/10.1088/1742-6596/1631/1/012141>

# **PŘÍLOHY**

# SEZNAM PŘÍLOH

## **Elektronické přílohy (dostupné na webové stránce této práce):**

- Příloha 1 Zdrojový kód – třída B-GRU-PTSAR (b\_gru\_ptsar.py)
- Příloha 2 Zdrojový kód – model pro COVID-19 (covid.py)
- Příloha 3 Dokumentace metod třídy B-GRU-PTSAR (documentation.pdf)
- Příloha 4 Tabulky s daty k predikcím (tables.zip)
- Příloha 5 Výsledný model (model.zip)
- Příloha 6 Vstupní data (data.zip)

## **Volné přílohy**

- Příloha 7 Poster ve formátu A2