Investigating the Evolution of the COVID-19 Pandemic in Germany Using Physics-Informed Neural Networks

Bachelor Thesis in Computer Science

submitted by
Phillip Rothenbeck
born February 22, 2002 in Eckernförde

written at
Computer Vision Group
Department of Mathematics and Computer Science
Friedrich-Schiller-Universität Jena

Supervisor:	Prof. Dr.-Ing. Joachim Denzler
Advisor:	Niklas Penzel, Sai Karthikeya Vemuri
Started:	May 1, 2024
Finished:	September 14, 2024

Eigenständigkeitserklärung

1. Hiermit versichere ich, dass ich die vorliegende Arbeit - bei einer Gruppenarbeit die von mir zu verantwortenden und entsprechend gekennzeichneten Teile - selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich trage die Verantwortung für die Qualität des Textes sowie die Auswahl aller Inhalte und habe sichergestellt, dass Informationen und Argumente mit geeigneten wissenschaftlichen Quellen belegt bzw. gestützt werden. Die aus fremden oder auch eigenen, älteren Quellen wörtlich oder sinngemäß übernommenen Textstellen, Gedankengänge, Konzepte, Grafiken etc. in meinen Ausführungen habe ich als solche eindeutig gekennzeichnet und mit vollständigen Verweisen auf die jeweilige Quelle versehen. Alle weiteren Inhalte dieser Arbeit ohne entsprechende Verweise stammen im urheberrechtlichen Sinn von mir.

2. Ich weiß, dass meine Eigenständigkeitserklärung sich auch auf nicht zitierfähige, generierende KI-Anwendungen (nachfolgend „generierende KI“) bezieht. Mir ist bewusst, dass die Verwendung von generierender KI unzulässig ist, sofern nicht deren Nutzung von der prüfenden Person ausdrücklich freigegeben wurde (Freigabeerklärung). Sofern eine Zulassung als Hilfsmittel erfolgt ist, versichere ich, dass ich mich generierender KI lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss deutlich überwiegt. Ich verantworte die Übernahme der von mir verwendeten maschinell generierten Passagen in meiner Arbeit vollumfänglich selbst. Für den Fall der Freigabe der Verwendung von generierender KI für die Erstellung der vorliegenden Arbeit wird eine Verwendung in einem gesonderten Anhang meiner Arbeit kenntlich gemacht. Dieser Anhang enthält eine Angabe oder eine detaillierte Dokumentation über die Verwendung generierender KI gemäß den Vorgaben in der Freigabeerklärung der prüfenden Person. Die Details zum Gebrauch generierender KI bei der Erstellung der vorliegenden Arbeit inklusive Art, Ziel und Umfang der Verwendung sowie die Art der Nachweispflicht habe ich der Freigabeerklärung der prüfenden Person entnommen.

3. Ich versichere des Weiteren, dass die vorliegende Arbeit bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt wurde oder in deutscher oder einer anderen Sprache als Veröffentlichung erschienen ist.

4. Mir ist bekannt, dass ein Verstoß gegen die vorbenannten Punkte prüfungsrechtliche Konsequenzen haben und insbesondere dazu führen kann, dass meine Prüfungsleistung als Täuschung und damit als mit „nicht bestanden“ bewertet werden kann. Bei mehrfachem oder schwerwiegendem Täuschungsversuch kann ich befristet oder sogar dauerhaft von der Erbringung weiterer Prüfungsleistungen in meinem Studiengang ausgeschlossen werden.

5. Die Richtlinien des Lehrstuhls für Examensarbeiten habe ich gelesen und anerkannt. Seitens des Verfassers/der Verfasserin bestehen keine Einwände, die vorliegende Examensarbeit für die öffentliche Benutzung zur Verfügung zu stellen.

Jena, den 14. September 2024 Phillip Rothenbeck

40 pages in total

Überblick

German version of the abstract.
Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Abstract

English version of the abstract.
Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Contents

Chapter 1

Introduction

1.1 Related work

Chapter 2

Theoretical Background

This chapter introduces the theoretical knowledge that forms the foundation of the work presented in this thesis. In Section 2.1 and Section 2.2, we talk about differential equations and the underlying theory. In these Sections both the explanations and the approach are strongly based on the book on analysis by Rudin [Rud07] and the book about ordinary differential equations by Tenenbaum and Pollard [TP85]. Subsequently, we employ this knowledge to examine various pandemic models in Section 2.3. Finally, we address the topic of neural networks with a focus on the multilayer perceptron in Section 2.4 and physics informed neural networks in Section 2.5.

2.1 Mathematical Modelling using Functions

To model a physical problem using mathematical tools, it is necessary to define a set of fundamental numbers or quantities upon which the subsequent calculations will be based. These sets may represent, for instance, a specific time interval or a distance. The term *domain* describes these fundamental sets of numbers or quantities [Rud07]. A *variable* is a changing entity living in a certain domain. In this thesis, we will focus on domains of real numbers in $\mathbb{R}$.

The mapping between variables enables the modeling of the process and depicts the semantics. We use functions in order to facilitate this mapping. Let $A, B \subset \mathbb{R}$ be to subsets of the real numbers, then we define a function as the mapping

$$f : A \to B. \tag{2.1}$$

In other words, the function f maps elements $x \in A$ to values $f(x) \in B$. A is the *domain* of f, while B is the *codomain* of f. Functions are capable of representing the state of a system as a value based on an input value from their domain. One

illustrative example is a function that maps a time point to the distance covered since a starting point. In this case, time serves as the domain, while the distance is the codomain.

2.2 Basics of Differential Equations

Often, the change of a system is more interesting than its current state. Functions are able to give us the latter, but only passively give information about the change of a system. The objective is to determine an effective method for calculating the change of a function across its domain. Let f be a function and $[a, b] \subset \mathbb{R}$ an interval of real numbers, the expression

$$m = \frac{f(b) - f(a)}{a - b} \tag{2.2}$$

gives the average rate of change. While the average rate of change is useful in many cases, the momentary rate of change is more accurate. To calculate this, we need to narrow down, the interval to an infinitesimal. For each $x \in [a, b]$ we calculate

$$\frac{df}{dx} = \lim_{t \to x} \frac{f(t) - f(x)}{t - x}, \tag{2.3}$$

if it exists. $\frac{df}{dx}$ is the *derivative*, or *differential equation*, it returns the momentary rate of change of f for each value x of f's domain. Repeating this process on $\frac{df}{dx}$ yields $\frac{d^2f}{dx^2}$, which is the function that calculates the rate of change of the rate of change and is called the second order derivative. Iterating this n times results in $\frac{d^nf}{dx^n}$, the derivative of the n'th order. Another method for obtaining a differential equation is to create it from the semantics of a problem. This method is useful if no basic function exists for a system. Differential equations find application in several areas such as engineering, physics, economics, epidemiology, and beyond.

Here insert definition of differential equations (take from books)

In the context of functions, it is possible to have multiple domains, meaning that function has more than one parameter. To illustrate, consider a function operating in two-dimensional space, wherein each parameter represents one axis or one that, employs with time and locations as inputs. The term *partial differential equations* (*PDE*'s) describes differential equations of such functions, which require a derivative for each of their domains. In contrast, *ordinary differential equations* (*ODE*'s) are the single derivatives for a function having only one domain. In this thesis, we only

need ODE's.

A *system of differential equations* is the name for a set of differential equations. The derivatives in a system of differential equations each have their own codomain, which is part of the problem, while they all share the same domain.

Tenenbaum and Pollard [TP85] provide many examples for ODE's, including the *Motion of a Particle Along a Straight Line*. Further, Newton's second law states that "the rate of change of the momentum of a body ($momentum = mass \cdot velocity$) is proportional to the resultant external force F acted upon it" [TP85]. Let m be the mass of the body in kilograms, v its velocity in meters per second and t the time in seconds. Then, Newton's second law translates mathematically to

$$F = m\frac{dv}{dt}. \tag{2.4}$$

It is evident that the acceleration, $a = \frac{dv}{dt}$, as the rate of change of the velocity is part of the equation. Additionally, the velocity of a body is the derivative of the distance traveled by that body. Based on these findings, we can rewrite the Equation (2.4) to

$$F = ma = m\frac{d^2s}{dt^2}. \tag{2.5}$$

This explanation of differential equations focuses on the aspects deemed crucial for this thesis and is not intended to be a complete explanation of the subject. To gain a better understanding of it, we recommend the books mentioned above [Rud07, TP85]. In the following section we describe the application of these principles in epidemiological models.

2.3 Epidemiological Models

Pandemics, like *COVID-19*, which has resulted in a significant number of fatalities. The question arises: How should we fight a pandemic correctly? Also, it is essential to study whether the employed countermeasures efficacious in combating the pandemic. Given the unfavorable public response to measures such as lockdowns, it is imperative to investigate that their efficacy remains commensurate with the costs incurred to those affected. In the event that alternative and novel technologies were in use, such as the mRNA vaccines in the context of COVID-19, it is needful to test

the effect and find the optimal variant. In order to shed light on the aforementioned events we need to develop a method to quantize the pandemic along with its course of progression.

The real world is a highly complex system, which presents a significant challenge attempting to describe it fully in a model. Therefore, the model must reduce the complexity while retaining the essential information. Furthermore, it must address the issue of limited data availability. For instance, during COVID-19 institutions such as the Robert Koch Institute (RKI)[1] were only able to collect data on infections and mortality cases. Consequently, we require a model that employs an abstraction of the real world to illustrate the events and relations that are pivotal to understanding the problem.

2.3.1 SIR Model

In 1927, Kermack and McKendrick [KM27] introduced the *SIR Model*, which subsequently became one of the most influential epidemiological models. This model enables the modeling of infections during epidemiological events such as pandemics. The book *Mathematical Models in Biology* [EK05] reiterates the model and serves as the foundation for the following explanation of SIR models.

The SIR model is capable of illustrating diseases, which are transferred through contact or proximity of an individual carrying the illness and a healthy individual. This is possible due to the distinction between infected beings who are carriers of the disease and the part of the population, which is susceptible to infection. In the model, the mentioned groups are capable to change, *e.g.*, healthy individuals becoming infected. The model assumes the size N of the population remains constant throughout the duration of the pandemic. The population N comprises three distinct groups: the *susceptible* group S, the *infectious* group I and the *removed* group R (hence SIR model). Let $\mathcal{T} = [t_0, t_f] \subseteq \mathbb{R}_{\geq 0}$ be the time span of the pandemic, then,

$$S : \mathcal{T} \to \mathbb{N}, \quad I : \mathcal{T} \to \mathbb{N}, \quad R : \mathcal{T} \to \mathbb{N}, \tag{2.6}$$

give the values of S, I and R at a certain point of time $t \in \mathcal{T}$. For S, I, R and N applies:

$$N = S + I + R. \tag{2.7}$$

[1] `https://www.rki.de/EN/Home/homepage_node.html`

The model makes another assumption by stating that recovered people are immune to the illness and infectious individual can not infect them. The individuals in the R group are either recovered or deceased, and thus unable to transmit or carry the disease. As visualized in the Figure 2.1 the individuals may transition between groups

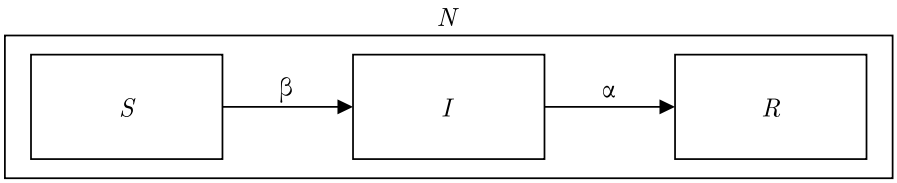


Figure 2.1: A visualization of the SIR model, illustrating N being split in the three groups S, I and R.

based on transition rates. The transmission rate β is responsible for individuals becoming infected, while the rate of removal or recovery rate α (also referred to as δ or ν, *e.g.*, [EK05, MPF23]) moves individuals from I to R.

We can describe this problem mathematically using a system of differential equations (see Section 2.2). Thus, Kermack and McKendrick [KM27] propose the following set of differential equations:

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI, \\ \frac{dI}{dt} &= \beta SI - \alpha I, \\ \frac{dR}{dt} &= \alpha I. \end{aligned} \tag{2.8}$$

This, according to Edelstein-Keshet, is based on the following assumption: "The rate of transmission of a microparasitic disease is proportional to the rate of encounter of susceptible and infective individuals modelled by the product (βSI)" [EK05]. The system shows the change in size of the groups per time unit due to infections, recoveries, and deaths.

The term βSI describes the rate of encounters of susceptible and infected individuals. This term is dependent on the size of S and I, thus Anderson and May [And91] propose a modified model:

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{SI}{N}, \\ \frac{dI}{dt} &= \beta \frac{SI}{N} - \alpha I, \\ \frac{dR}{dt} &= \alpha I. \end{aligned} \tag{2.9}$$

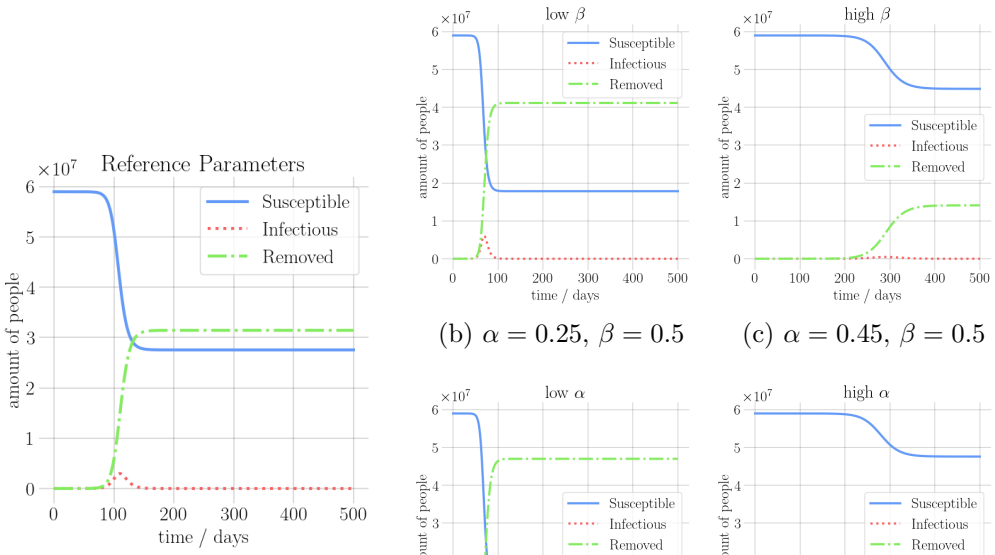
In which βSI gets normalized by N, which is more correct in a real world aspect [And91].

The initial phase of a pandemic is characterized by the infection of a small number of individuals, while the majority of the population remains susceptible. The infectious group has not yet infected any individuals thus neither recovery nor mortality is possible. Let $I_0 \in \mathbb{N}$ be the number of infected individuals at the beginning of the disease. Then,

$$\begin{aligned} S(0) &= N - I_0, \\ I(0) &= I_0, \\ R(0) &= 0, \end{aligned} \tag{2.10}$$

describes the initial configuration of a system in which a disease has just emerged.

In the SIR model the temporal occurrence and the height of the peak (or peaks) of the infectious group are of paramount importance for understanding the dynamics of a pandemic. A low peak occurring at a late point in time indicates that the disease is unable to keep pace with the rate of recovery, resulting in its demise before it can exert a significant influence on the population. In contrast, an early and high peak means that the disease is rapidly transmitted through the population, with a significant proportion of individuals having been infected. Figure 2.1 illustrates the impact of modifying either β or α while simulating a pandemic using a model such as Equation (2.9). It is evident that both the transmission rate β and the recovery rate α influence the height and time of the peak of I. When the number of infections exceeds the number of recoveries, the peak of I will occur early and will be high. On the other hand, if recoveries occur at a faster rate than new infections the peak



(a) $\alpha = 0.35$, $\beta = 0.5$

(b) $\alpha = 0.25$, $\beta = 0.5$

(c) $\alpha = 0.45$, $\beta = 0.5$

(d) $\alpha = 0.35$, $\beta = 0.4$

(e) $\alpha = 0.35$, $\beta = 0.6$

Figure 2.2: Synthetic data, using Equation (2.9) and $N = 7.9 \cdot 10^6$, $I_0 = 10$ with different sets of parameters.

will occur later and will be low. This means, that it is crucial to know both β and α to be able to simulate a pandemic using the SIR model.

The SIR model makes a number of assumptions that are intended to reduce the model's overall complexity while simultaneously increasing its divergence from actual reality. One such assumption is that the size of the population, N, remains constant. This depiction is not an accurate representation of the actual relations observed in the real world, as the size of a population is subject to a number of factors that can contribute to change. The population is increased by the occurrence of births and decreased by the occurrence of deaths. There are different reasons for mortality, including the natural aging process or the development of other diseases. Other examples are the absence of the possibility for individuals to be susceptible again, after having recovered, or the possibility for the transition rates to change due to new variants or the implementation of new countermeasures. We address this latter option in the next Section 2.3.2.

2.3.2 Reduced SIR Model and the Reproduction Number

The Section 2.3.1 presents the classical SIR model. The model comprises two parameters β and α, which describe the course of a pandemic over its duration. This is beneficial when examining the overall pandemic; however, in the real world, disease behavior is dynamic, and the values of the parameters β and α change at each time point. The reason for this is due to events such as the implementation of countermeasures that reduce the contact between the infectious and susceptible individuals, the emergence of a new variant of the disease that increases its infectivity or deadliness, or the administration of a vaccination that provides previously susceptible individuals with immunity without ever being infectious. To address this Millevoi *et al.* [MPF23] introduce a model that simultaneously reduces the size of the system of differential equations and solves the problem of time scaling at hand.

First, they alter the definition of β and α to be dependent on the time interval $\mathcal{T} = [t_0, t_f] \subseteq \mathbb{R}_{\geq 0}$,

$$\beta : \mathcal{T} \to \mathbb{R}_{\geq 0}, \quad \alpha : \mathcal{T} \to \mathbb{R}_{\geq 0}. \tag{2.11}$$

Another crucial element is $D(t) = \frac{1}{\alpha(t)}$, which represents the initial time span an infected individual requires to recuperate. Subsequently, at the initial time point t_0, the *reproduction number*,

$$\mathcal{R}_0 = \beta(t_0)D(t_0) = \frac{\beta(t_0)}{\alpha(t_0)}, \tag{2.12}$$

represents the number of susceptible individuals, that one infectious individual infects at the onset of the pandemic.In light of the effects of β and α (see Section 2.3.1), $\mathcal{R}_0 > 1$ indicates that the pandemic is emerging. In this scenario α is relatively low due to the limited number of infections resulting from $I(t_0) << S(t_0)$. When $\mathcal{R}_0 < 1$, the disease is spreading rapidly across the population, with an increase in I occurring at a high rate. Nevertheless, $\mathcal{R}_0$ does not cover the entire time span. For this reason, Millevoi *et al.* [MPF23] introduce $\mathcal{R}_t$ which has the same interpretation as $\mathcal{R}_0$, with the exception that $\mathcal{R}_t$ is dependent on time. The definition of the time-dependent reproduction number on the time interval $\mathcal{T}$ with the population size N,

$$\mathcal{R}_t = \frac{\beta(t)}{\alpha(t)} \cdot \frac{S(t)}{N} \tag{2.13}$$

includes the rates of change for information about the spread of the disease and information of the decrease of the ratio of susceptible individuals in the population. In contrast to β and α, $\mathcal{R}_t$ is not a parameter but a state variable in the model and enabling the following reduction of the SIR model.

Equation (2.7) allows for the calculation of the value of the group R using S and I, with the term $R(t) = N - S(t) - I(t)$. Thus,

$$\begin{aligned} \frac{dS}{dt} &= \alpha(\mathcal{R}_t - 1)I(t), \\ \frac{dI}{dt} &= -\alpha\mathcal{R}_t I(t), \end{aligned} \tag{2.14}$$

is the reduction of Equation (2.8) on the time interval $\mathcal{T}$ using this characteristic and the reproduction number $\mathcal{R}_t$(see Equation (2.13)). Another issue that Millevoi *et al.* [MPF23] seek to address is the extensive range of values that the SIR groups can assume, spanning from 0 to 10^7. Accordingly, they initially scale the time interval $\mathcal{T}$ using its borders to calculate the scaled time $t_s = \frac{t-t_0}{t_f-t_0} \in [0, 1]$. Subsequently, they calculate the scaled groups,

$$S_s(t_s) = \frac{S(t)}{C}, \quad I_s(t_s) = \frac{I(t)}{C}, \quad R_s(t_s) = \frac{R(t)}{C}, \tag{2.15}$$

using a large constant scaling factor $C \in \mathbb{N}$. Applying this to the variable I, results in,

$$\frac{dI_s}{dt_s} = \alpha(t_f - t_0)(\mathcal{R}_t - 1)I_s(t_s), \tag{2.16}$$

a further reduced version of Equation (2.8) results in a more streamlined and efficient process, as it entails the elimination of a parameter(β) and two state variables (S and R), while adding the state variable $\mathcal{R}_t$. This is a crucial aspect for the automated resolution of such differential equation systems, as we describe in Section 2.4.

2.4 Multilayer Perceptron

In Section 2.2 we show the importance of differential equations to systems, being able to show the change of it dependent on a certain parameter of the parameter. In Section 2.3 we show specific applications for differential equations in an epidemiological context. Now, the last point is to solve these equations. For this problem, there are multiple methods to reach this goal one of them is the *Multilayer Percep-*

tron (MLP) [HSW89]. In the following we briefly tackle the structure, training and usage of these *neural networks* using, for which we use the book *Deep Learning* by Goodfellow *et al.* [GBC16] as a base for our explanations.

The goal is to be able to approximate any function f^* that is for instance mathematical function or a mapping of an input vector to a class or category. Let $\boldsymbol{x}$ be the input vector and $\boldsymbol{y}$ the label, class or result, then,

$$\boldsymbol{y} = f^*(\boldsymbol{x}), \tag{2.17}$$

is the function to approximate. In the year 1958, Rosenblatt [Ros58] proposed the perceptron modeling the concept of a neuron in a neuroscientific sense. The perceptron takes in the input vector $\boldsymbol{x}$ performs anoperation and produces a scalar result. This model optimizes its parameters θ to be able to calculate $\boldsymbol{y} = f(\boldsymbol{x}; \theta)$ as correct as possible. As Minsky and Papert [MP72] show, the perceptron on its own is able to approximate only a class of functions. Thus, the need for an expansion of the perceptron.

As Goodfellow *et al.* go on, the solution for this is to split f into a chain structure of $f(\boldsymbol{x}) = f^{(3)}(f^{(2)}(f^{(1)}(\boldsymbol{x})))$. This transforms a perceptron, which has an input and output layer into a multilayer perceptron. Each sub-function $f^{(n)}$ is represented in the structure of an MLP as a *layer*, which are each build of a multitude of *units* (also *neurons*) each of which are doing the same vector-to-scalar calculation as the perceptron does. Each scalar, is then given to a nonlinear activation function. The layers are staggered in the neural network, with each being connected to its neighbor, in the way as illustrated in Figure 2.3. The input vector $\boldsymbol{x}$ is given to each unit of the first layer $f^{(1)}$, which results are then given to the units of the second layer $f^{(2)}$, and so on. The last layer is called the *output layer*. All layers in between the first and the output layers are called *hidden layers*. Through the alternating structure of linear and nonlinear calculation MLP's are able to approximate any kind of function. As Hornik *et al.* [HSW89] shows, MLP's are universal approximators.

The process of optimizing the parameters θ is called *learning*. Here, we define a metric for the quality of the results, of our neural network. This metric is called a loss function

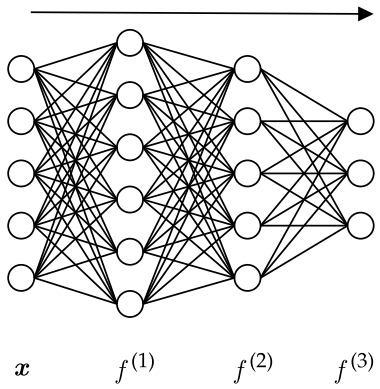


Figure 2.3: A visualization of the SIR model, illustrating N being split in the three groups S, I and R.

2.5 Physics Informed Neural Networks

2.5.1 Disease Informed Neural Networks

Chapter 3

Methods

3.1 Data Preprocessing

3.1.1 RKI Data

3.1.2 Recovery Queue

3.2 PINN for the SIR Model

3.3 PINN for the reduced SIR Model

Chapter 4

Experiments

4.1 SIR Model

4.1.1 Setup

4.1.2 Results

4.2 Reduced SIR Model

4.2.1 Setup

4.2.2 Results

Chapter 5

Conclusions

5.1 Further Work

Bibliography

[And91] Anderson, Robert M. Roy Malcolm; May M. Roy Malcolm; May: *Infectious diseases of humans : dynamics and control*. Oxford University Press, 1991

[EK05] Edelstein-Keshet, Leah: *Mathematical Models in Biology*. Society for Industrial and Applied Mathematics, 2005

[GBC16] Goodfellow, Ian ; Bengio, Yoshua ; Courville, Aaron: *Deep Learning*. MIT Press, 2016. – http://www.deeplearningbook.org

[HSW89] Hornik, Kurt ; Stinchcombe, Maxwell ; White, Halbert: Multilayer feedforward networks are universal approximators. In: *Neural Networks* 2 (1989), Januar, Nr. 5, S. 359–366. http://dx.doi.org/10.1016/0893-6080(89)90020-8. – DOI 10.1016/0893–6080(89)90020–8. – ISSN 0893–6080

[KM27] Kermack, William O. ; McKendrick, A. G.: A contribution to the mathematical theory of epidemics. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115 (1927), August, Nr. 772, S. 700–721. http://dx.doi.org/10.1098/rspa.1927.0118. – DOI 10.1098/rspa.1927.0118. – ISSN 2053–9150

[MP72] Minsky, Marvin ; Papert, Seymour A.: *Perceptrons*. 2. print. with corr. Cambridge/Mass. [u.a.] : The MIT Press, 1972. – ISBN 9780262630221. – Literaturangaben

[MPF23] Millevoi, Caterina ; Pasetto, Damiano ; Ferronato, Massimiliano: A Physics-Informed Neural Network approach for compartmental epidemiological models. (2023). http://dx.doi.org/10.48550/ARXIV.2311.09944. – DOI 10.48550/ARXIV.2311.09944

[Ros58] Rosenblatt, F.: The perceptron: A probabilistic model for information storage and organization in the brain. In: *Psychological Review* 65 (1958), Nr. 6, S. 386–408. http://dx.doi.org/10.1037/h0042519. – DOI 10.1037/h0042519. – ISSN 0033–295X

[Rud07] Rudin, Walter: *Analysis*. Oldenbourg Wissenschaftsverlag GmbH, 2007

[TP85] Tenenbaum, Morris ; Pollard, Harry: *Ordinary Differential Equations*. Harper and Row, Publishers, Inc., 1985

List of Figures

List of Tables