

Hier den Titel ohne Formatierung einfügen

STUDIENARBEIT

für die Prüfung zum
Bachelor of Engineering
des Studienganges Informationstechnik
an der
Dualen Hochschule Baden-Württemberg Karlsruhe
von

Alexander Bierenstiel, André Schmitt, Dominik Schmitt

Abgabedatum 14. Mai 2018

Bearbeitungszeitraum	900 Stunden
Matrikelnummer	xxxxxxxxx,yyyyyyy, 7191584
Kurs	TINF15B3
Ausbildungsfirma	Firmennamen
	Firmenort
Betreuer der Ausbildungsfirma	Namehier
Gutachter der Studienakademie	Prof. Dr. Jürgen Vollmer

Erklärung

Ich versichere hiermit, dass ich meine Studienarbeit mit dem Thema: „Titelhiernochmal“ selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Ort Datum

Unterschrift

Sofern von der Ausbildungsstätte ein Sperrvermerk gewünscht wird, ist folgende Formulierung zu verwenden:

Sperrvermerk

Der Inhalt dieser Arbeit darf weder als Ganzes noch in Auszügen Personen außerhalb des Prüfungsprozesses und des Evaluationsverfahrens zugänglich gemacht werden, sofern keine anders lautende Genehmigung der Ausbildungsstätte vorliegt.

Zusammenfassung

Dieses Abstract besser schreiben und eventuell eine englische Übersetzung anfertigen

Die vorliegende Studienarbeit befasst sich mit dem Thema der deutschen Bahn und ihrer Verspätungen. Es soll die von der Bahn zu Verfügung gestellten API genutzt werden, um Daten zu sammeln. Anhand dieser Daten soll ein neuronales Netz modelliert werden, welches genutzt werden kann, um Verspätungen und Abhängigkeiten im Schienenverkehr zu erkennen und vorherzusagen.

Inhaltsverzeichnis

1	Einleitung	6
1.1	Usefulcmds	6
1.2	Stand der Technik	8
1.3	Ziel der Studienarbeit	8
2	Grundlagen	9
2.1	Einarbeitung	9
2.2	Planung	9
2.3	Data Mining	9
2.4	Aufbereitung von Daten	9
2.5	Neuronalen Netzen an simplen Beispielen erklärt	9
2.6	Literaturhinweise und Empfehlungen	10
3	Umsetzung	11
3.1	Programmierung des Data Miners	11
3.2	Programmierung der Automatischen Datenverarbeitung	11
3.3	Programmieren der Modellstruktur des neuronalen netzes	11
3.4	Vermeidung von Overfitting und Anpassungen um die Genauigkeit zu erhöhen	11
3.5	Anlernen des Netzes	12
3.6	Verifizieren des Netzes	12
3.7	Prediction anhand des Netzes	12
3.8	Auswertung und Fehlerbehandlung	12
4	Schlussfolgerung	13
4.1	Rückblick	13
4.2	Fazit	13
4.3	Ausblick	13
	Anhang	14
	Literaturverzeichnis	14
	Liste der ToDo's	15

Abbildungsverzeichnis

1.1	SomeCaption	6
1.2	Schöne darfstellung	7

Tabellenverzeichnis

Liste der Quellcodeausschnitte

1.1	Pseudo Code	7
-----	-----------------------	---

Abkürzungsverzeichnis

RFC	Request for Comments.....	7
------------	---------------------------	---

Kapitel 1

Einleitung

1.1 Usefullcmds

Hier werden nützliche Commands aufgezeigt.

aaa

ccc

„Alles klar hier könnte der Text stehen“

[KNUTH 1984]

Ohne Zitats hier bla bla [DANTE 2010]



Abbildung 1.1: SomeCaption

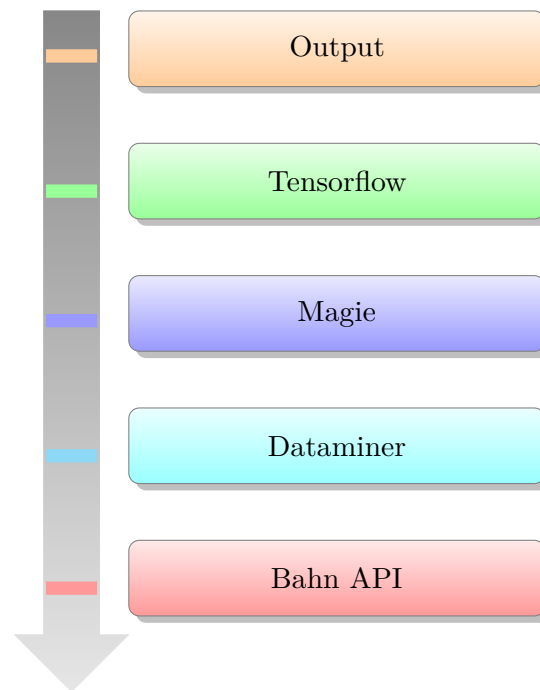


Abbildung 1.2: Schöne darfstellung

```

1 function getAllValues($value) {
2     // Comment
3     private static $foo = "bar";
4     $bar = TRUE;
5     return $foo;
6 }

```

Quellcode 1.1: Pseudo Code

io Hallo welt io ist eine Wort ab sofort wenn ich das sage Hallo welt io ist eine Wort ab
sofort wenn ich das sage Hallo welt io ist eine Wort ab sofort wenn ich das sage

bar foobar

Request for Comments (RFC) 4227¹ siehe hier nochmal RFC
(Siehe Quellcode 1.1)

Tensorflow, Graphviz, Pencil v3 (ist super zum schnellen erstellen von Abläufen o.ä.),
Python 3.5, 3.6

¹Siehe [KNUTH 1984]

1.2 Stand der Technik

Hier etwas zum Stand der Technik schreiben, neuronale Netze, Tensorflow, KI, Data-mining, OpenData, etc

1.3 Ziel der Studienarbeit

Hier das Ziel aus der Anmeldung schön definieren und klar abgrenzen was Ziel und was optional nice to have ist.

Kapitel 2

Grundlagen

2.1 Einarbeitung

Wo beginnen wir und was müssen wir alles beachten

2.2 Planung

Zeitliche einteilung, beachten 5. Semester ist weniger Zeit, Hauptteil wird im 6. Semester passieren

2.3 Data Mining

Data Mining Einführung und dessen Bedeutung für das Projekt

2.4 Aufbereitung von Daten

Wie werden Daten aufbereitet, vorbereitet für das neuronale Netz, welche dinge gibt es zu beachten (DATENTYPEN!)

2.5 Neuronale Netze an simplen Beispielen erklärt

Kleine Einleitung an einem Simplen Beispiel, Linear Regression oder so. Wieso wird sowas brauchen und weshalb es von relevanz ist.

2.6 Literaturhinweise und Empfehlungen

Weiterführende Literatur sollte bis zum Abschluss erwähnt werden, verwendete Quellen zum Einlesen in neuronale Netze und gute Erklärungen, event. Zitate auch benutzen. Diese Autoren sind xsehr wichtig für dieses Projekt und sollte auch genannt werden.

Kapitel 3

Umsetzung

3.1 Programmierung des Data Miners

Wie funktioniert der Data miner, wie erkennt er Fehler, was passiert bei Fehlern

3.2 Programmierung der Automatischen Datenverarbeitung

Wie kommen die Datensatz aus der DB zum neuronalen Netzwerk, wie wird die Formatierung vorgenommen

3.3 Programmieren der Modellstruktur des neuronalen netzes

Wie ist das neuronale Netz strukturiert, welche neuronen verwenden wir und weshalb, wie finden wir das beste Netz

3.4 Vermeidung von Overfitting und Anpassungen um die Genauigkeit zu erhöhen

Overfitting ist häufig ein Problem wieoe erkennt man es und wie kann man overfitting vermeiden.

3.5 Anlernen des Netzes

Welche Datensätze werden zum Anlernen verwendet, weshalb ist es wichtig nie alle zu nehmen im Bezug auf Test, Predict und welche verhältnisse sind bei uns sinnvoll anzusetzen

3.6 Verifizieren des Netzes

Testen des neuronalen Netzes, verifikation der Genauigkeit und deren Steigerung durch Training oder anpassungen des Netzes

3.7 Prediction anhand des Netzes

Vorhersagen aus Daten treffen und anschauen wie gut sie sind, wo gibt es Probleme, welche Probleme treten auf.

3.8 Auswertung und Fehlerbehandlung

Was passiert im Fehlerfall, wie erkennt man Fehler, müssen wir Fehler erkennen oder sind Fehler egal", wie stellen wir ein GUI bereit, um anderen Menschen die Ergebnisse zu testen, genauere Statistiken zu Zügen je nach Strecke, Uhrzeit etc., vllt Visuelle Darstellung awie bei travic oder mit eigenen Heatmaps bzw. Openstreetmap.

Kapitel 4

Schlussfolgerung

4.1 Rückblick

Was ist geschehen, was würden wir anderst machen, was war das wichtigste Rückschauend

4.2 Fazit

Ergebnis der Studienarbeit, was war gut, was war schlecht, hat alles so geklappt, wo gab es Probleme, wie wurden diese gelöst (kurz und knapp zusammengefasst).

4.3 Ausblick




















Wie geht es weiter, könnte es weiter gehen, was sollte verbessert werden, wo befinden sich Schwachstellen, event. ungelöste Probleme

Literatur

DANTE [2010]. *Webseite der Deutschsprachige Anwendervereinigung TeX e.V.* <http://www.dante.de> [siehe S. 6].

KNUTH, Donald E. [1984]. *The T_EXbook*. Addison-Wesley [siehe S. 6, 7].

Liste der ToDo's

	Dieses Abstract besser schreiben und eventuell eine englische Übersetzung anfertigen	1
	Hier werden nützliche Commands aufgezeigt.	6
	aaa	6
	ccc	6
	Hier etwas zum Stand der Technik schreiben, neuronale Netze, Tensorflow, KI, Datamining, OpenData, etc	8
	Hier das Ziel aus der Anmeldung schön definieren und klar abgrenzen was Ziel und was optional nice to have ist.	8
	Wo beginnen wir und was müssen wir alles beachten	9
	Zeitliche einteilung, beachten 5. Semester ist weniger Zeit, Hauptteil wird im 6. Semester passieren	9
	Data Mining Einführung und dessen Bedeutung für das Projekt	9
	Wie werden Daten aufbereitet, vorbereitet für das neuronale Netz, welche dinge gibt es zu beachten (DATENTYPEN!)	9
	Kleine Einleitung an einem Simplen Beispiel, Linear Regression oder so. Wieso wird sowas brauchen und weshalb es von relevanz ist.	9
	Weiterführende Literatur sollte bis zum Abschluss erwähnt werden, verwendete Quellen zum Einlesen in neuronale Netze und gute Erklärungen, event. Zitate auch benutzen. Diese Autoren sind xsehr wichtig für dieses Projekt und sollte auch genannt werden.	10
	Wie funktioniert der Data miner, wie erkennt er Fehler, was passiert bei Fehlern	11
	Wie kommen die Datensatz aus der DB zum neuronalen Netzwerk, wie wird die Formatierung vorgenommen	11
	Wie ist das neuronale Netz strukturiert, welche neuronen verwenden wir und weshalb, wie finden wir das beste Netz	11
	Overfitting ist häufig ein Problem wioe erkennt man es und wie kann man overfitting vermeiden.	11
	Welche Datensätze werden zum Anlernen verwendet, weshalb ist es wichtig nie alle zu nehmen im Bezug auf Test, Predict und welche verhältnisse sind bei uns sinnvoll anzusetzen	12
	Testen des neuronalen Netzes, verifikation der Genauigkeit und deren Steigerung durch Training oder anpassungen des Netzes	12
	Vorhersagen aus Daten treffen und anschauen wie gut sie sind, wo gibt es Probleme, welche Probleme treten auf.	12

■ Was passiert im Fehlerfall, wie erkennt man Fehler, müssen wir Fehler erkennen oder sind Fehler egal", wie stellen wir ein GUI bereit, um anderen Menschen die Ergebnisse zu testen, genauere Statistiken zu Zügen je nach Strecke, Uhrzeit etc., vllt Visuelle Darstellung awie bei travic oder mit eigenen Heatmaps bzw. Openstreetmap.	12
■ Was ist geschehen, was würden wir anderst machen, was war das wichtigste Rückschauend	13
■ Ergebnis der Studienarbeit, was war gut, was war schlecht, hat alles so geklappt, wo gab es Probleme, wie wurden diese gelöst (kurz und knapp zusammengefasst.	13
■ Wie geht es weiter, könnte es weiter gehen, was sollte verbessert werden, wo befinden sich Schwachstellen, event. ungelöste Probleme	13