Introduction
○○○

State of the Art
○○○○○○

Prototype
○○○

Evaluation
○○

IT-based Automatic Text Summarization with the Use of Textgeneration Methods

Löhr Tim

Technische Hochschule Nürnberg Georg Simon OHM

Bachelor Thesis — Business Information Systems and Management

May 13, 2020

TECHNISCHE HOCHSCHULE NÜRNBERG
GEORG SIMON OHM

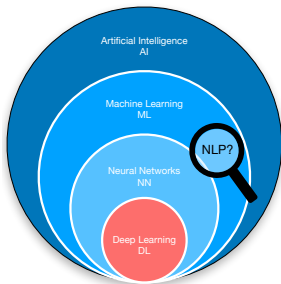## Overview

Introduction

Introduction
○●○

State of the Art
○○○○○○

Prototype
○○○

Evaluation
○○

Clarify the Keywords

Noticeably almost annually Artificial Intelligence is finding more and more its way into businesses.

Famous phrases for Advertisement

1. *Our product is powered now by AI!*
2. *We now use Deep Learning for a better performance!*

In conclusion: Deep Learning is a technique making use of Neural networks. Those are methods of Machine Learning, which is itself just an application of the entire AI ecosystem.

Introduction
○○●

State of the Art
○○○○○○

Prototype
○○○

Evaluation
○○

Classify my Thesis in this Ecosystem

NLP

Natural Language Processing (NLP) deals and manipulates data in form of our language to gain new information from it or perform other related tasks such as Text Summarization.

Where is NLP located?

Depending on what is performed, Natural Language Processing is either just an ML application, or like in my case, it can even make use of Neural Networks with Deep Learning methods to generate better results!

Introduction
○○○

State of the Art
●○○○○○

Prototype
○○○

Evaluation
○○

State of the Art

Introduction
ooo

State of the Art
o●oooo

Prototype
ooo

Evaluation
oo

General Methology

Using data analysis and show how it can be used to improve the
marketing of Airbnb in Seattle according to each policy of the
marketing-mix.

Marketing

The 4 P's of the Marketing-Mix

1. Product
2. Price
3. Promotion
4. Place

Definition Marketing

Marketing is about the firm's effort to address customer needs and expectations, which influences the demands made by the customers on the product and need to be fulfilled by the product.

Introduction
000

State of the Art
000●00

Prototype
000

Evaluation
00

How to use Data Analysis to improve the Airbnb Marketing?

- **Descriptive analysis:** Reviews, locations, review and price correlation, details of listings and price correlation
- **Descriptive analysis:** Predict the number of customer
- **Optimization:** Optimize the booking of listings
- **Adaptive learning:** Learn from the results generated and combine results to give out suggestion in marketing campaigns may hold by Airbnb

Introduction
○○○

State of the Art
○○○○●○

Prototype
○○○

Evaluation
○○

Reviews comparison

We can see a correlation between the amount of reviews and the accommodation counts.

Introduction
000

State of the Art
000000●

Prototype
000

Evaluation
00
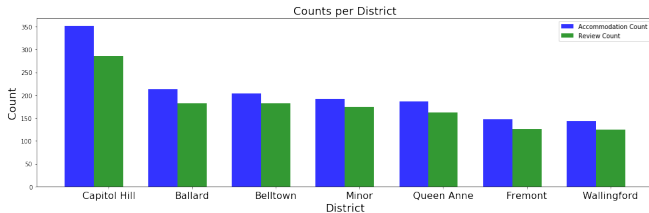
Linear Regression

- Use a Linear function and estimate its parameters
- There are different approaches to estimate the parameters
- The accuracy of a model can be compared with different loss functions

The fitted line can mathematically described as:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \qquad (1)$$

Introduction
000

State of the Art
000000

Prototype
●00

Evaluation
00

Prototype

Introduction
000

State of the Art
000000

Prototype
0●0

Evaluation
00

Multinomial Naive Bayes

- Assumes conditionally independent classes
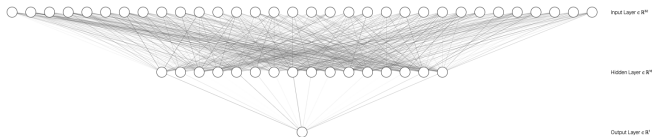- Probability of observing features f_1 through f_n, given some class c

$$p(f_1, ..., f_n|c) = \prod_{i=1}^{n} p(f_i|c)$$

(2)

This means that when I want to use a Naive Bayes model to classify a new example, the posterior probability is much simpler to work with:

$$p(c|f_1, ..., f_n) \propto p(c)p(f_1|c)...p(f_n|c)$$

(3)

Introduction
000

State of the Art
000000

Prototype
00●

Evaluation
00

LSTM Neural Network

- Long Short Term Memory Neural Network is based on the Recurrent Neural Network
- It is very well suited for timeseries analysis like the prediction of price
- Futher explanation exceeds this presentation

Introduction
000

State of the Art
000000

Prototype
000

Evaluation
●○

Evaluation

Introduction
000

State of the Art
000000

Prototype
000

Evaluation
0●

The team around Josh Keating reached a mean absolute error between 32\$ to 35\$

- Our network design was way to simple
- We didn't include the important neighbourhood feature
- Our result with our LSTM neural network was 64.04\$ (MSA)
- Nevertheless, our best model was the linear regression with a MAE of 43.89\$