

Abstract

The subject of this thesis is the domain of *prediction* using supervised learning. More precisely, we approach the problem of *text prediction*. Text prediction is a subfield of Natural Language Processing (NLP) which is an important research field in Artificial Intelligence , concerned with the interaction between computers and human languages.

The thesis is structured in four chapters. In the first chapter , we begin with briefly describing artificial neural networks, followed by a more in depth presentation of the recurrent neural networks and the Long Short-Term Memory (LSTM) networks and how they are good models for analyzing sequences of data. The second chapter presents the theory on Hidden Markov Models (HMM) , taking a closer look on the usual problems that these models are used for. The next chapter highlights the current state of the art of text prediction and completion, offering different perspectives and results. The final chapter describes our software application for text prediction, giving details about its design and implementation.

An original aspect of the thesis consists in presenting, in an original manner, the current state of the art in the field of supervised learning and text prediction. Another original contribution consists in designing, implementing and documenting the application presented in the last chapter which demonstrates the applicability of the Hidden Markov Models in the problem of text prediction.

This work is the result of my own activity. I have neither given nor received unauthorized assistance on this work.

Teodora Ioana NASTE

Cluj-Napoca
23.06.2019