

# Chatbot Epsilon



The Sympathetic Encyclopaedia

*Jesper, Marcus, Niklas och Rasmus*

# Introduction

## Aim

- Create a simple chatbot using an artificial network (ANN) to analyze text responses typed by a user.

## Two groups

- Sentiment Analysis
- Conversation

## Chat flow

# Conversation

## Methods

- RNN
- Cosine Similarity
- Intent Classification

## Datasets

- Wikipedia
- Reddit
- SQuAD 2.0



S  
2.0

# Conversation

## Final Model

- List of Sports
- Lemmatization
- Tokenization
- Cosine Similarity
  - Generate response
  - Subject selection
    - Sports on Wikipedia

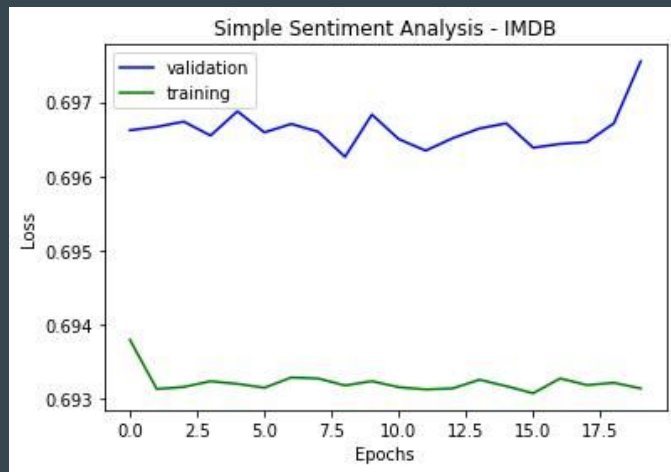
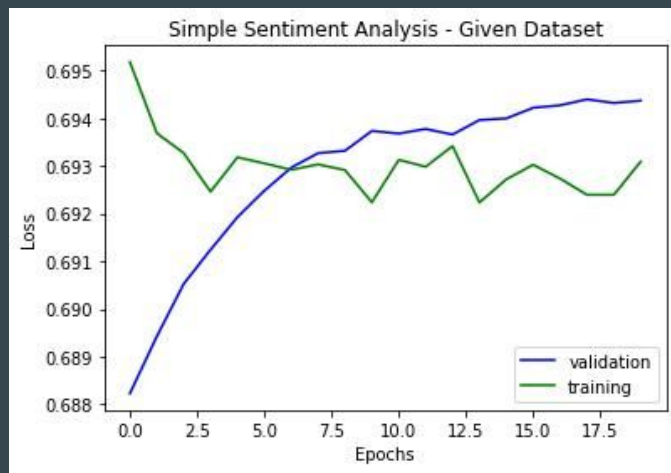
## Improvements

- Intent classification

# Sentiment Analysis

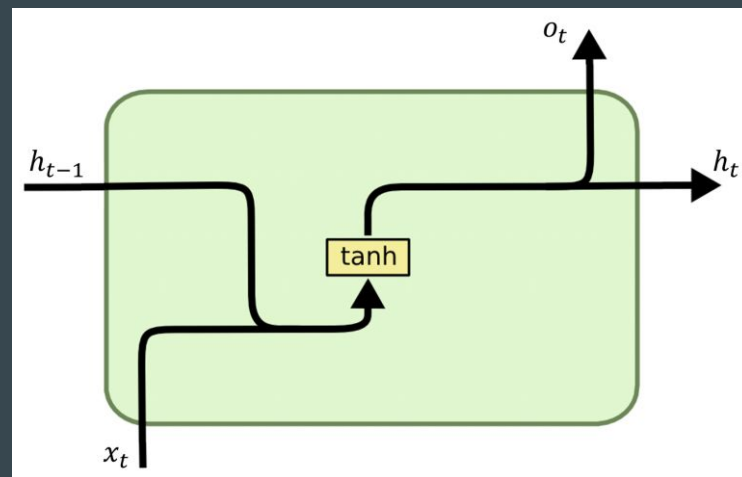
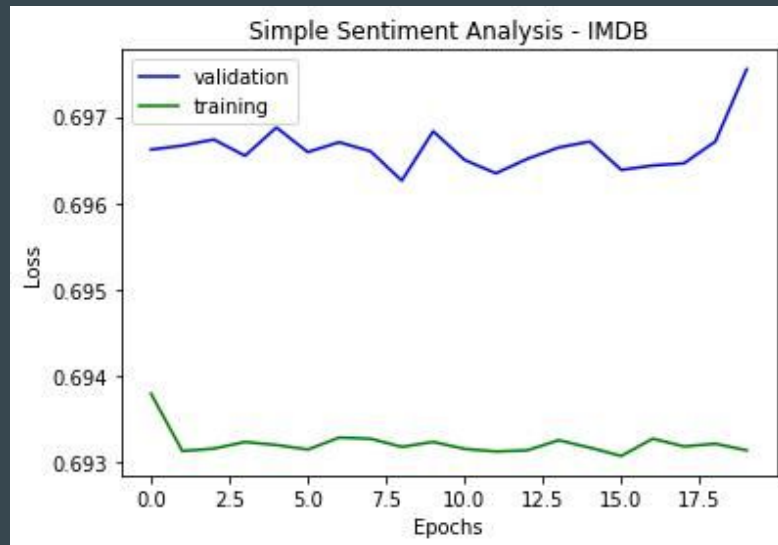
## 5 Approaches

- Simple (RNN)
- Upgraded (LSTM)
- FastText
- Convolutional
- BertGRU



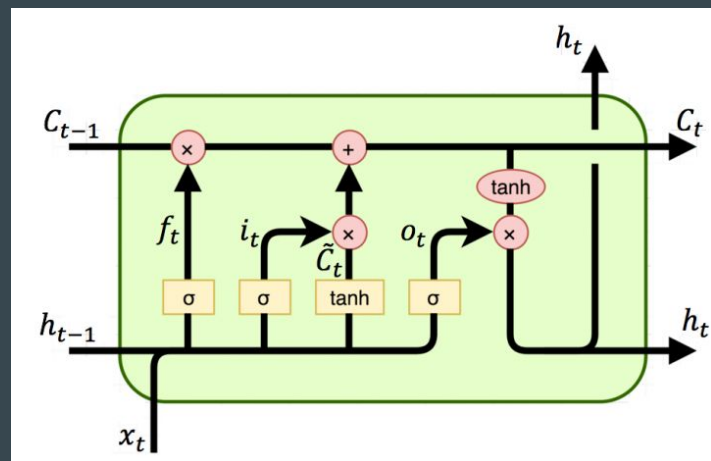
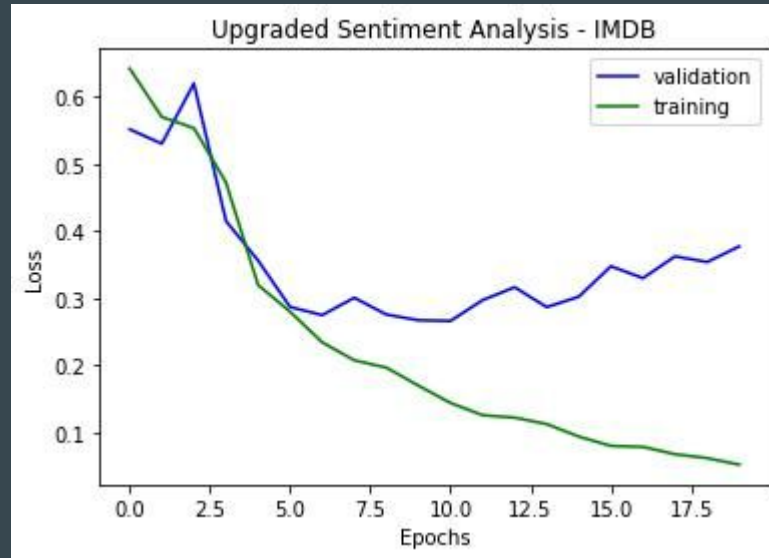
# Simple (RNN)

- Randomly initialized parameter
- SGD optimizer
- BCE loss function
- 2.8 Million parameters
- About 50% test accuracy
- Training time per epoch: 16s



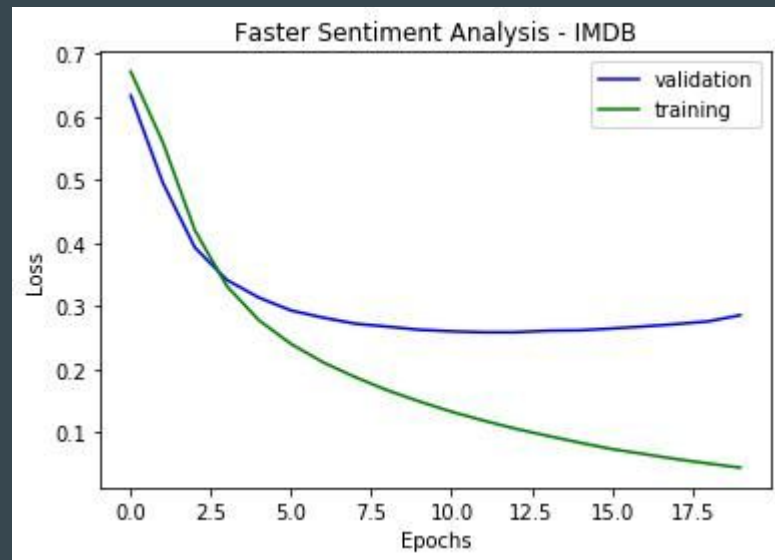
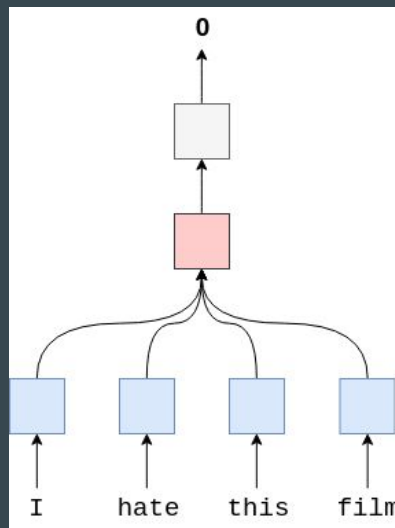
# Upgraded (LSTM)

- Bidirectional, multilayer LSTM
- GloVe 100d6B pre-trained embeddings
- Dropout of 50%
- Adam optimizer
- BCE loss function
- 4,8 Million trainable Parameters
- About 84% test accuracy
- Training time per epoch: 27s



# FastText

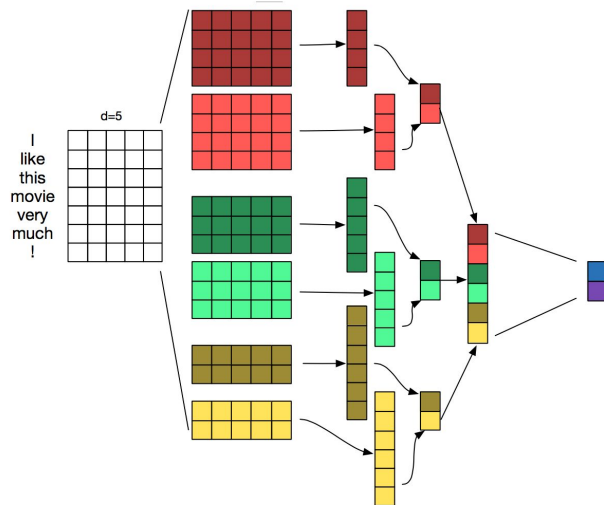
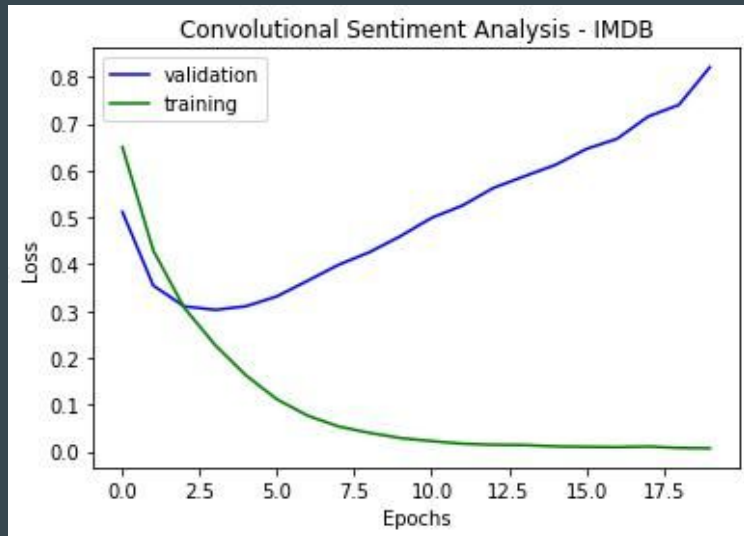
- Bi-grams
- GloVe 100d6B
- 2d Average pooling layer
- Adam optimizer
- BCE loss function
- 2,5 Million trainable Parameters
- About 85% test accuracy
- Training time per epoch: 5s





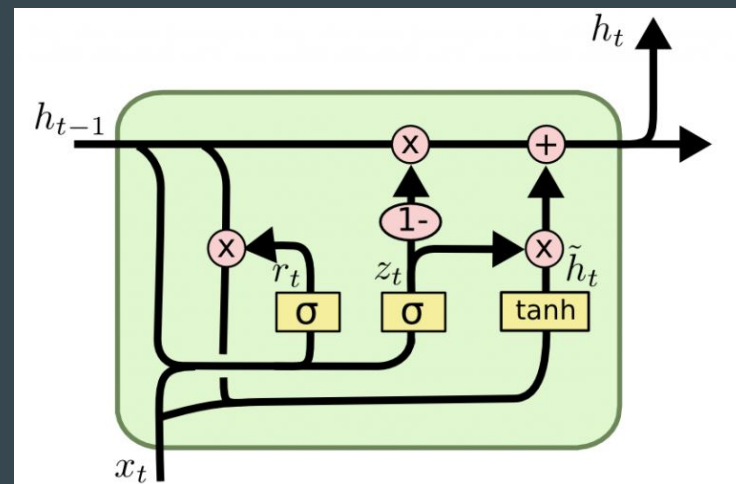
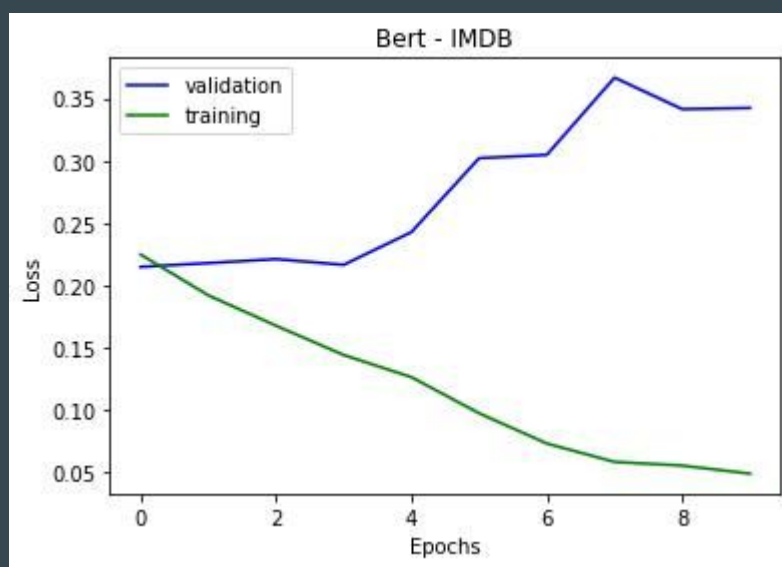
# Convolutional

- GloVe 100d6B
- Filter sizes of 3, 4 and 5 (100 of each)
- Max pooling and ReLU
- Dropout 50%
- 2,6 Million trainable parameters
- About 86% test accuracy
- Training time per epoch: 11s



# BertGRU

- Freeze all but GRU and Output
- Dropout of 25%
- Adam optimizer
- BCE loss function
- 2,5 Million trainable parameters
- About 92% test accuracy
- Training time per epoch: 8 min



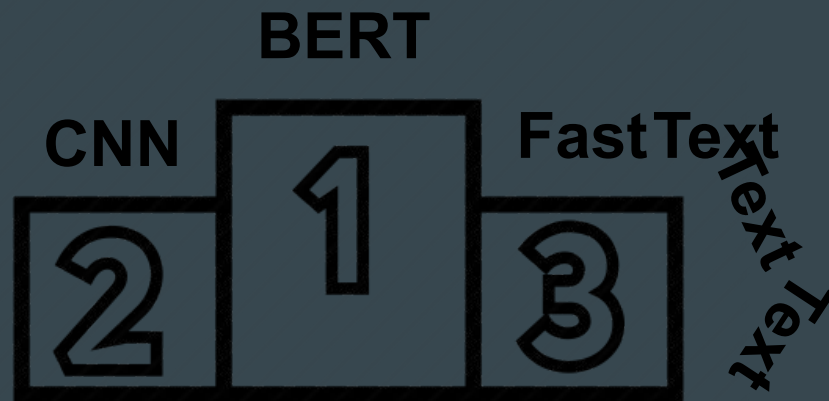
# Table of changes

Model	RNN	LSTM	FastText	CNN	BertGRU
Embedding	Random	GloVe 100d6B	GloVe 100d6B	GloVe 100d6B	Bert
Optimizer	SGD	Adam	Adam	Adam	Adam
Loss function	BCE	BCE	BCE	BCE	BCE
Trainable parameters	2.8M	4.8M	<u>2.5M</u>	2.6M	<u>2.5M</u>
Training time per epoch	16s	27s	<u>5s</u>	11s	8min
Test Accuracy	~50%	~84%	85%	86%	<u>92%</u>

# Sentiment - Conclusion

The best results were obtained using Bert.

However, for most of the project the CNN network was used, since it was fast to train with acceptable result.



# Demo



The Sympathetic Encyclopaedia