

Introduction

Today's newspapers have the power to shape one's entire perspective on the world. The idea is to create, train, adjust and improve a model so that it is fit to analyze various Swiss newspapers and to determine which papers are written with the most negative and positive attitude.

Project pipeline

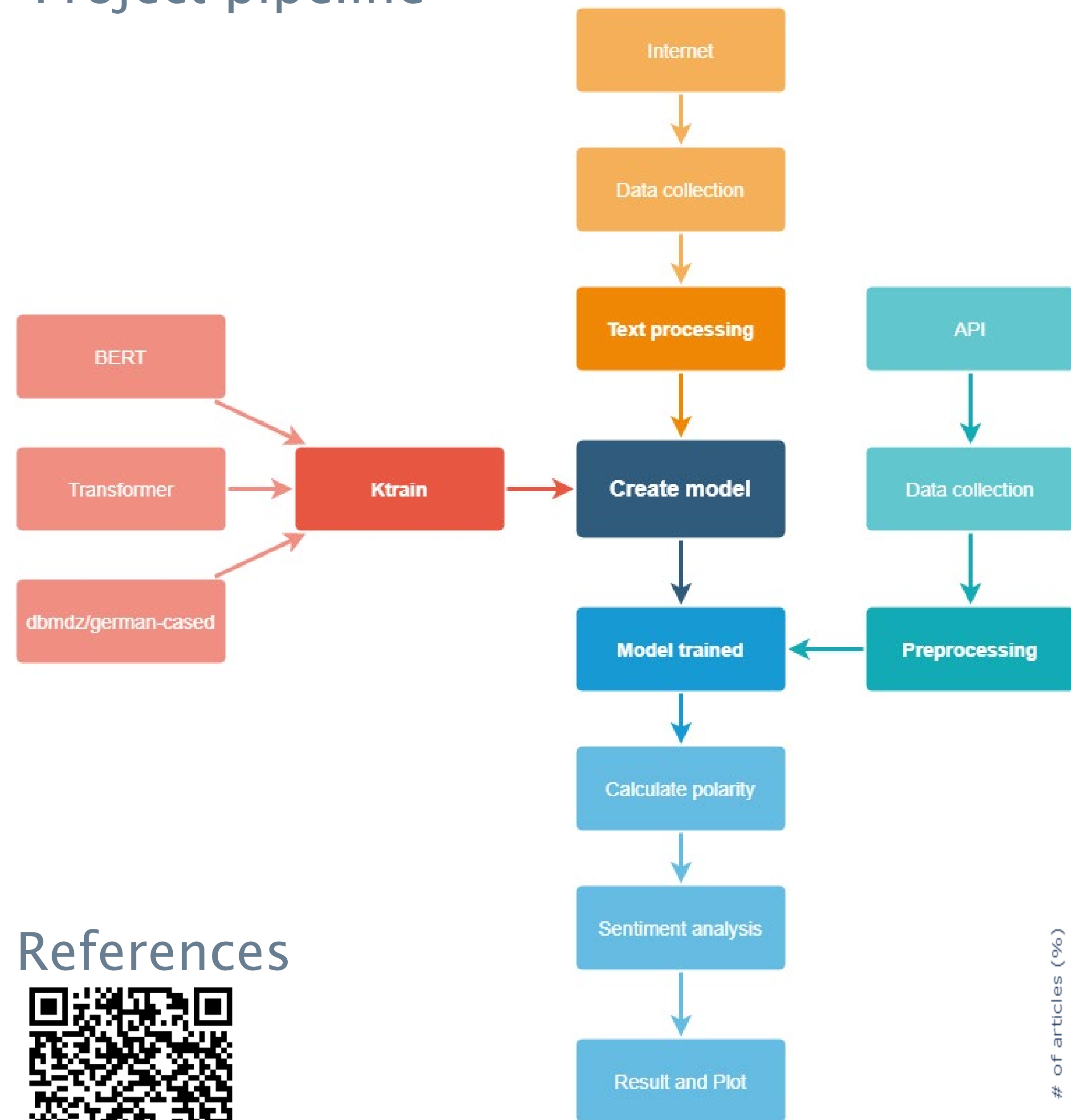


Fig 1: project pipeline

References



Methods

1. Model setup
 1. BERT, HuggingFace, Ktrain
 2. Filmstarts database
2. Newspaper database setup
 1. Data collection from APIs
 2. Preprocessing APIs data
 3. Automation
3. Sentiment Analysis
 1. Feature engineering
 2. Polarity calculation
 3. Plotting
 4. Analyze results

Model

In the process of the work, I created several models, each model being an improvement of the previous one. The first two models are based on the English language, then I trained the later models for the German language.

Model	Dataset	Accuracy
Ktrain1	Hotel Reviews (EN)	85%
Ktrain1 2ver	Hotel Reviews (EN)	96%
Ktrain2 auto	Filmstarts (DE)	90%
Ktrain3 manual	Filmstarts (DE)	93%

Sentiment Analysis

From the graphs I could see that most of the articles (in German) that we read in Switzerland have negative sentiments. Percentage is 60% NEG/40% POS(Fig2). In Fig3 is a comparison between 2 newspaper.

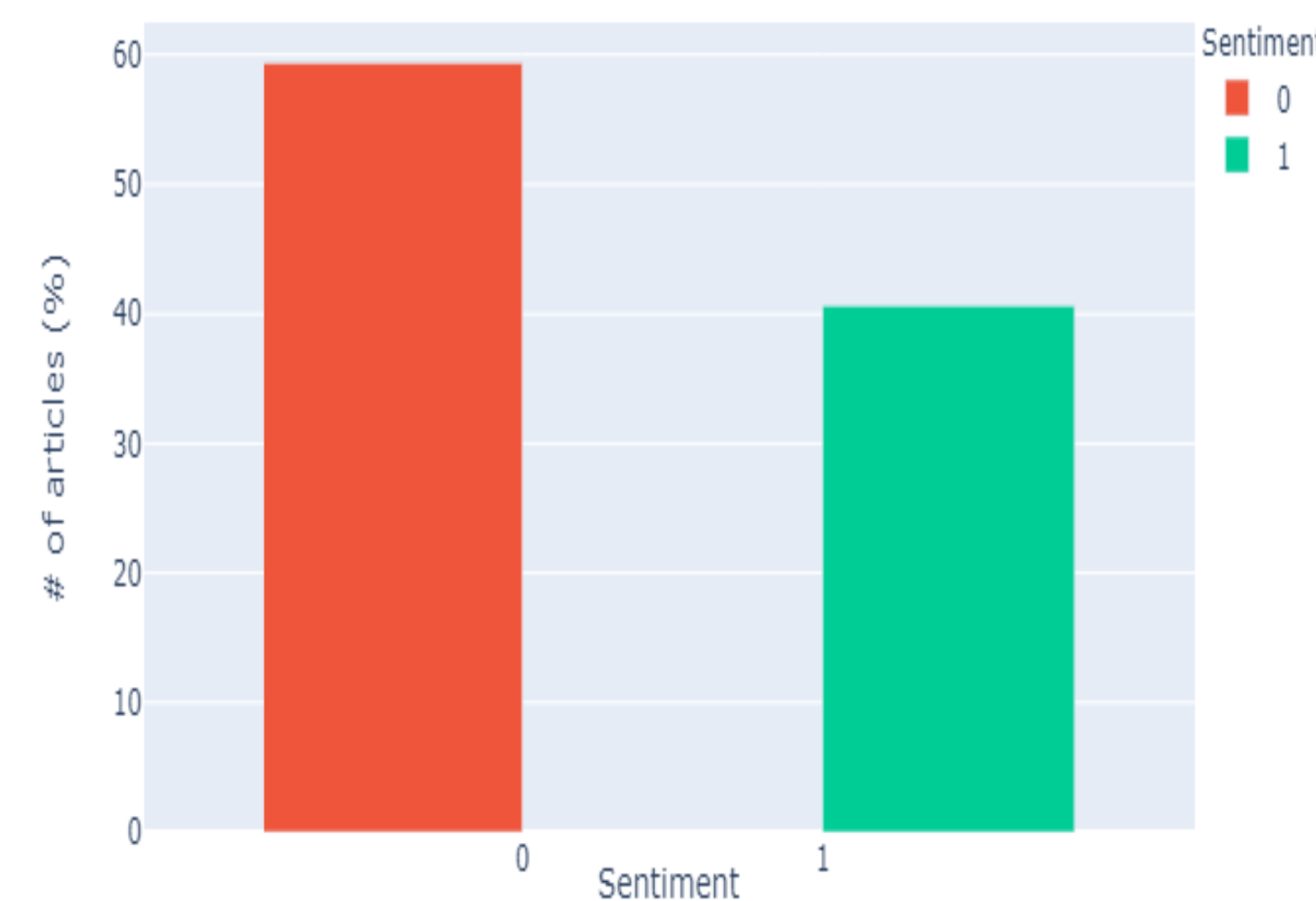
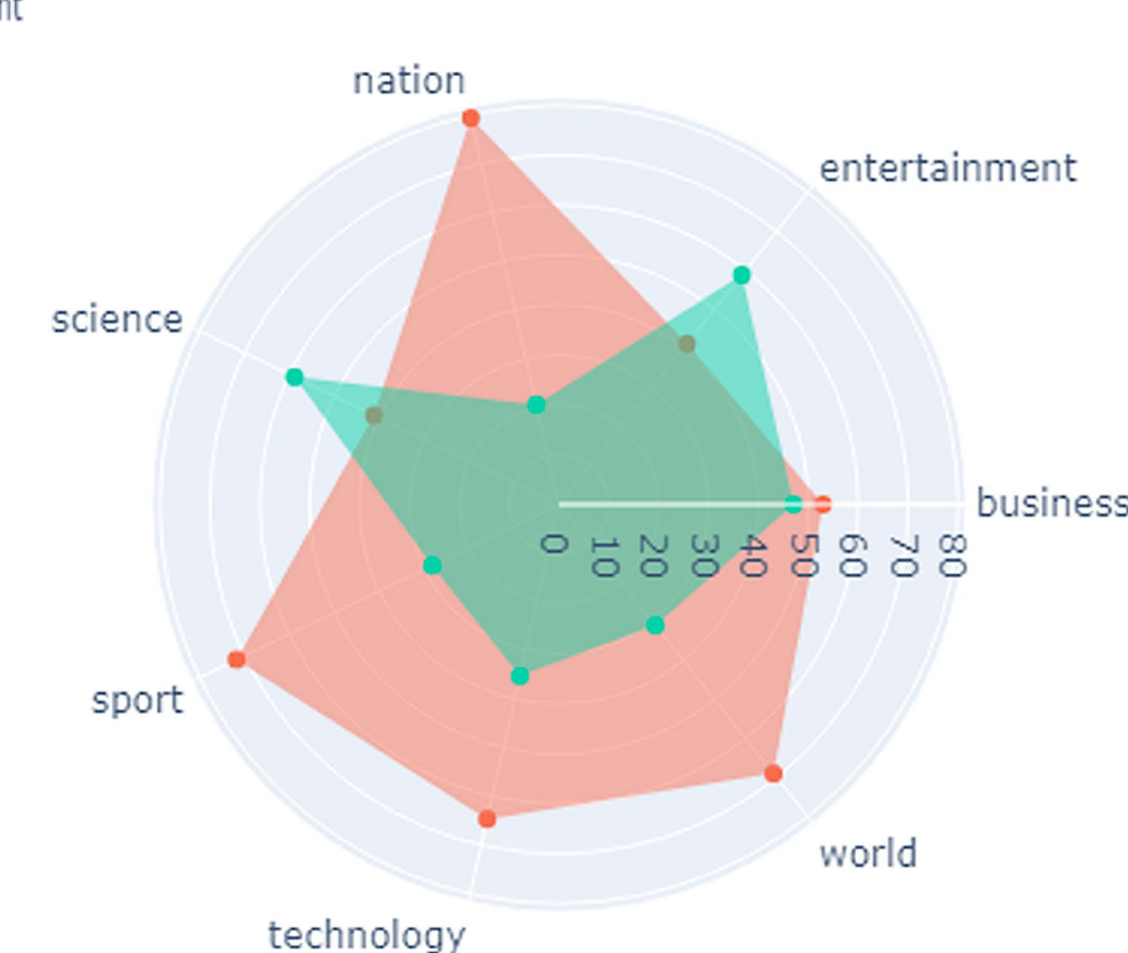


Fig 2: % sentiment

Spider Comparison: 20 minuten



Spider Comparison: blick

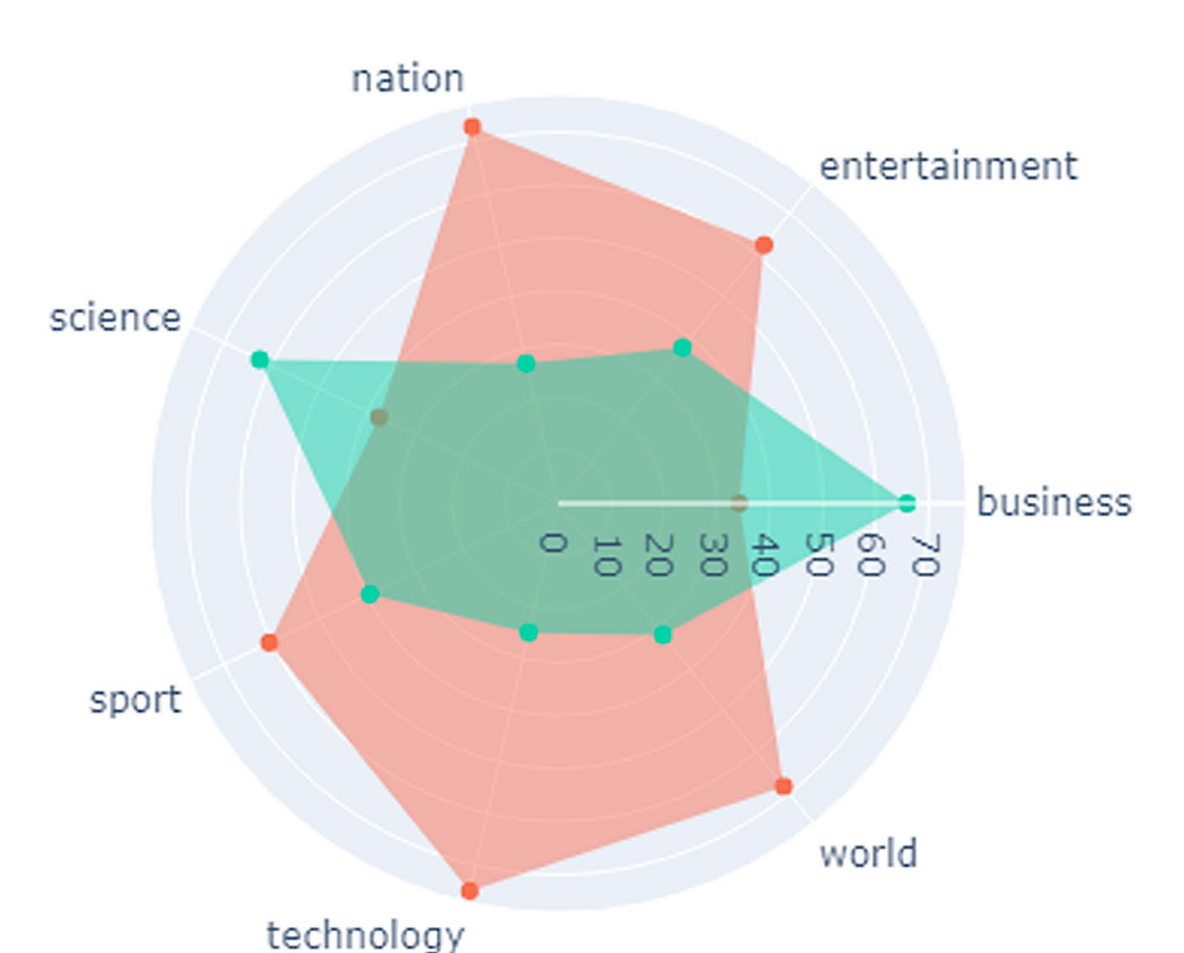


Fig 3: spider sentiment on «20 minuten» and «blick»

Conclusion

This work could very well be part of studying a psychological thesis on the masses, to see how the polarity of the news can influence our life.