Text Mining South Park

Dominik Nerger (i6146759)

June 2, 2017

Contents

1 Introduction

This report is part of the second assignment for the course Computer Vision and focusses on face recognition using Principal Component Analysis and k-nearest neighbors classification.

2 Data set
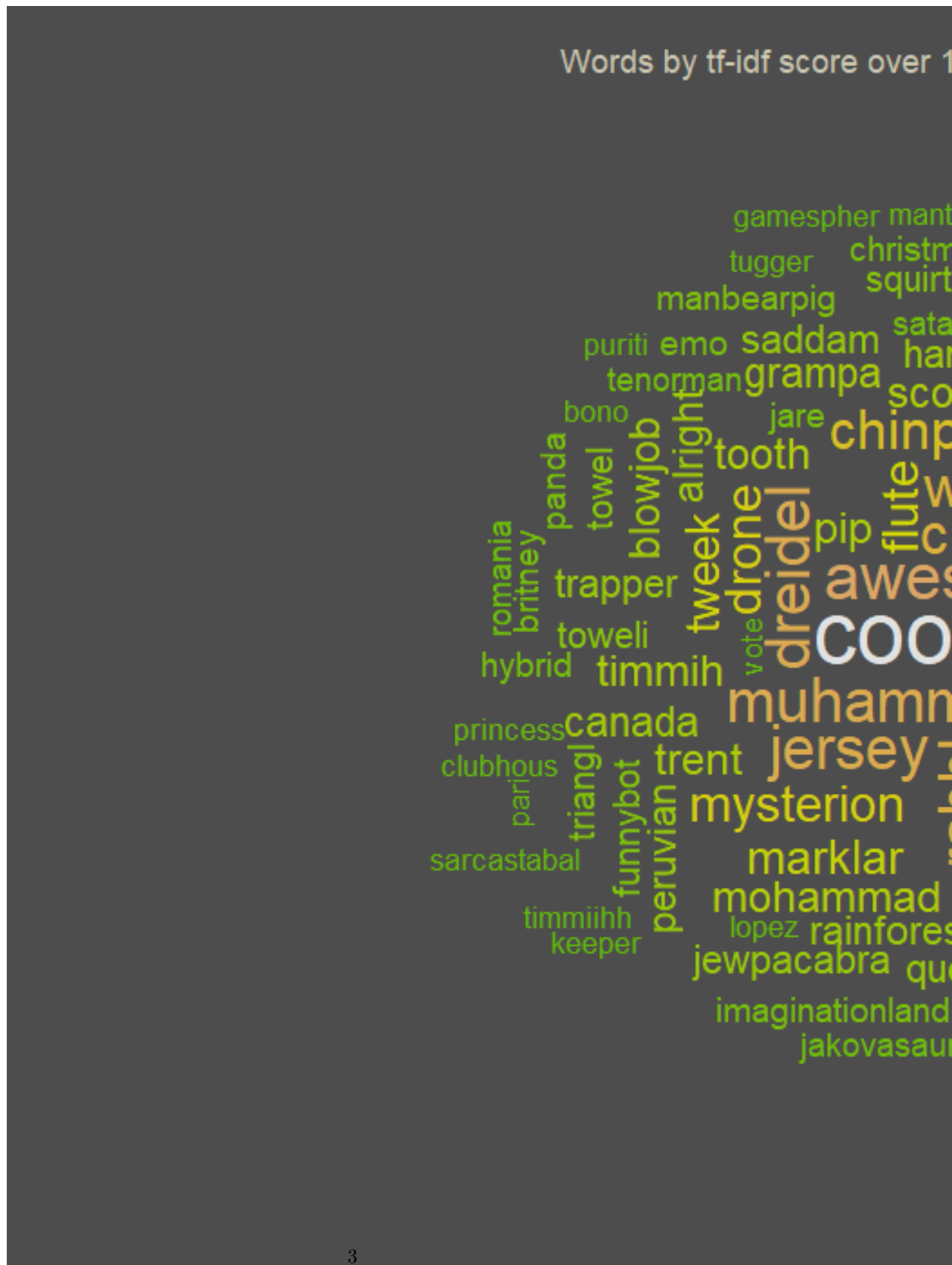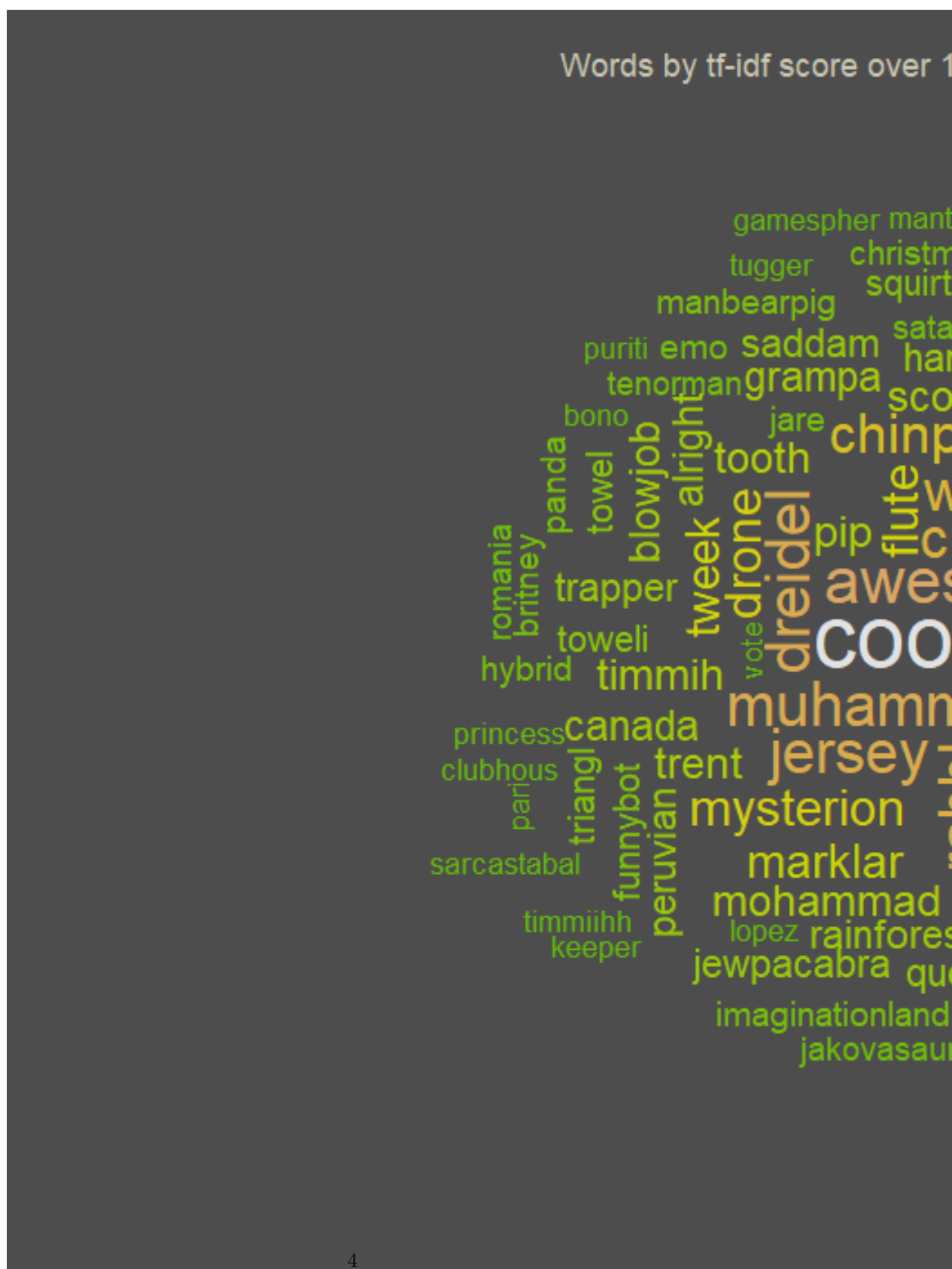
Figure 1: General wordcloud, containing terms by frequency

Figure 2: General wordcloud, containing terms by TF-IDF score

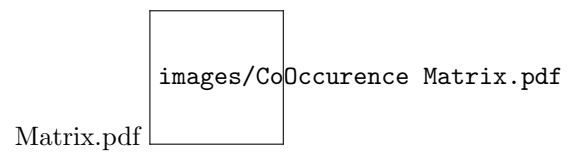Figure 3: Wordcloud containing ngrams (n=4) by frequency

Figure 4: Wordcloud containing ngrams (n=4) by frequency

3 Pre-processing

4 Implementation and results

5 Conclusion

In conclusion,

6 Future work