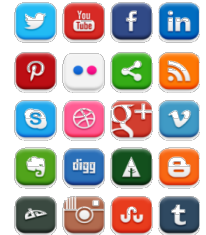




## Text Analytics using Supervised Text Classification

### Description

Text classification has become a popular solution to many problems brought forward by social media platforms such as Twitter. In a nutshell, supervised text classification extracts relevant information from manually classified documents and learns a model from the extracted information. The growing text analytics market heavily relies on text classification to offer services such as sentiment analysis, document categorization, scientific discovery etc. With the help of DKPro-TC, our framework for supervised text classification, you will study novel text analysis tasks and collect new insights into artificial intelligence.



<http://www.graphicsfuel.com/2013/03/popular-social-media-icons-psd-png>

### Possible Tasks

- Automatic Quality Assessment of Forum posts or Blog comments
- Vandalism Detection in Wikipedia and other collaboratively created resources
- Develop a Demo Web Interface for applying previously trained models to unseen data
- Bring your own data: Discover the potential of machine learning on textual data of your own choice

**Note:** Your task and its scope are subject to discussion and will be adapted to the targeted thesis type.

### References

- Daxenberger, J., Ferschke, O., Gurevych, I., & Zesch, T. (2014). DKPro TC: A Java-based Framework for Supervised Learning Experiments on Textual Data. In *Proceedings of the 52nd Annual Meeting of the ACL*.
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34(1), 1–47.
- Yano, T., Cohen, W. W., & Smith, N.A. (2009). Predicting response to political blog posts with topic models. In *Proceedings of the 2009 Annual Conference of the NAACL*. (pp. 477–485).

### Contact

Analysis      ■■■■□

Programming    ■■■■□

Literature      ■■■□□

Prof. Dr. Iryna Gurevych

Dr. Johannes Daxenberger

[thesis@ukp.informatik.tu-darmstadt.de](mailto:thesis@ukp.informatik.tu-darmstadt.de)