Sebastian Ruder

sebastianruder.com • ruder.sebastian@gmail.com
+353 89 237 9192 • 9 New Bride Street • Dublin, Ireland

Experience

- **AYLIEN** **Dublin, Ireland**
  NLP Engineer *10/2015 – Present*
  - I leverage my research in sentiment analysis to generate business value for AYLIEN, e.g. I implemented a system for aspect-based sentiment analysis of hotel reviews that can be adapted to other domains.
- **INSIGHT Centre for Data Analytics** **Galway, Ireland**
  Research Assistant *10/2015 – Present*
- **IBM** **Munich, Germany**
  Extreme Blue Intern, Watson *08/2015 – 09/2015*
  - Design and implementation of text analysis ML components applied to customer data of leading German insurance company VKB; automatically identifies semantics and sentiments of incoming e-mails, e.g. complaints.
- **Microsoft** **Dublin, Ireland**
  Linguistic Engineering Intern, XBox *02/2015 – 06/2015*
  - Contributed to developing an ML system for analyzing linguistic complexity in C#; performed feature analysis.
  - Created proof of concept and implemented terminology validation using morphology generation.
  - Evangelized customer sentiment analysis efforts, drove cross-team collaboration.
- **The OpenCog Foundation** opencog.org
  Google Summer of Code Intern *Summer 2014*
  - Enabled system to make common-sense inferences and deductions ("Socrates is mortal", etc.)
  - Applied inference using probabilistic logic networks on the output of a relationship extractor.
  - Documented and extended undocumented temporal inference Python code.
- **Lingenio GmbH** **Heidelberg, Germany**
  Software Engineering Intern *Spring 2014*
  - Created a converter from TBX in Lingenio native format and vice versa.
  - Integrated TBX term bases in Dictionary Server; localization: Jinja2, Flask-Babel; configuration: lighttpd.
- **SAP** **Walldorf, Germany**
  Working Student, Development University *02/2013 – 02/2014*
  - Created content for internal programming and Design Thinking courses.
  - Automated reporting processes, e.g. reduced expenditure of work for monthly training report by > 75%, i.e. from 8 hours to 2 hour using advanced Excel / VBA scripts.
- **TEMIS** **Heidelberg, Germany**
  Freelancing Developer *02/2013 – 10/2013*
  - Created a cosine metric-based word sense disambiguation system leveraging tens of GB of text extracted from Wikipedia and DBpedia in Python; achieved performance comparable to the state-of-the-art.

Education

- **National University of Ireland** **Galway, Ireland**
  College of Engineering and Informatics, Ph.D. Natural Language Processing *10/2015 – Present*
  - Research in applying ML and deep learning techniques to cross-lingual sentiment analysis in different domains as well as aspect- and entity-based. I'm interested in developing new algorithms and data sources to improve sentiment analysis systems and adapting them to other languages by leveraging existing monolingual / bilingual data as well as generating new data.

- **Ruprecht-Karls-Universität Heidelberg** **Heidelberg, Germany**
  Institute of Computational Linguistics, B.A. Computational Linguistics, English Linguistics *10/2012 – 09/2015*
  - Final grade: 1.0 (German scale), i.e. A+; thesis: *Construction and Analysis of an Emotion Proposition Store*
  - Relevant courses: Statistics, Algorithms and Data Structures, Machine Learning, Formal Syntax & Semantics
  - Relevant online courses: Machine Learning (Stanford), AI (MIT), Into to Algorithms (Berkeley), Deep Learning for NLP (Stanford), Deep Learning (Oxford)
- **Trinity College** **Dublin, Ireland**
  School of Computer Science and Statistics, Computer Science and Language *09/2014 – 01/2015*
  - Semester abroad
  - relevant courses: AI, Fuzzy Logic, High-Tech Entrepreneurship

Awards

- Scholarship of the Irish Research Council *10/2015 – Present*
- Scholarship of the German state *04/2014 – 09/2015*
- Microsoft Certified Professional (Programming in C#) *06/2015*
- Best Delegate award in various Model United Nations conferences *11/2012 – 01/2014*

Projects

- **Travel activity recommendation** *05/2015*
  Implemented a ML system that recommends travel activities to a user based on an analysis of her social media posts using a classifier trained on 400k Tweets pertaining to 50+ travel activities (85% accuracy) as part of a hackathon project.
- **Emotion detection** *05/2015 – 07/2015*
  Pattern-based extraction, storage, and distributional analysis of emotion propositions in the Gigaword corpus.
- **Coreference resolution** *04/2014 – 07/2014*
  Implemented the state-of-the-art Stanford sieve-based approach to coreference resolution in the modular BART system and extended it to achieve state-of-the-art performance in German.

Languages and Technologies

Programming Languages: Python, Java, C#, R, C, LaTeX, Prolog, JavaScript, SPARQL
Technologies: SciPy, NumPy, Keras, TensorFlow, NLTK, CoreNLP, MALLET, Weka, UNIX, Git
Natural Languages: Fluent in German and English, advanced in French and Spanish, beginner in Portuguese and Latin
Open Source Contributions: The OpenCog Foundation