Melanie Tosik

New York, NY

tosik@nyu.edu
melanietosik.com

Education

New York University, New York, NY   expected 05/2019

M.S., Computer Science

University of Potsdam, Potsdam, Germany   10/2011 – 09/2015

B.S., Computational Linguistics (GPA: 3.6)

Thesis: "String to Semantic Graph Alignment"

Professional Experience

Data Science Consultant   07/2017 – 12/2017
Underdog.io
New York, NY
Automated internal processes for resume parsing and candidate-company matching

Research Engineer – NLP   10/2015 – 06/2017
WayBlazer
Austin, TX
Designed and implemented a proprietary question answering system based on state-of-the-art techniques
for natural language understanding and ontology extraction

Research Intern   06/2014 – 08/2014
Textkernel
Amsterdam, The Netherlands
Improved resume parsing for multilingual documents by integrating word embeddings as input features
for a probabilistic sequence labeling model based on conditional random fields

Undergraduate Research Assistant   07/2013 – 04/2014
Applied Computational Linguistics Discourse Research Lab
University of Potsdam
Co-developed a Python tool for automatic coreference resolution within project D1 (Linguistic Database
for Information Structure: Annotation and Retrieval) of DFG funded Collaborative Research Center 632

Publications

Melanie Tosik, Carsten L. Hansen, Gerard Goossen, and Mihai Rotaru. "Word Embeddings vs
Word Types for Sequence Labeling: the Curious Case of CV Parsing." In *Workshop on Vector
Space Modeling for NLP, NAACL 2015*, Denver.

- Proficient in Python, Java, C++, R, Git, AWS, NoSQL, Elasticsearch, and related technologies

- Experienced with common NLP/ML toolkits and frameworks, including spaCy, SciPy, NumPy, CoreNLP, PyTorch, IBM Bluemix, etc.

- Familiar with ML algorithms for clustering, classification, sequence labeling, and deep learning

- Versed in knowledge base construction and entity linking using RDF, SPARQL, and DBpedia

Poster and presentations

Melanie Tosik. "Generating personalized travel recommendations from natural language queries." *Data Day Seattle*, Seattle, 2016.

Carsten L. Hansen, **Melanie Tosik**, Gerard Goossen, Chao Li, Lena Bayeva, Florence Berbain, and Mihai Rotaru. "How to Get the Best Word Vectors for Resume Parsing." In *SNN Adaptive Intelligence / Symposium: Machine Learning 2015*, Nijmegen, 2015.

Teaching Experience

Teaching Assistant   Spring 2018
    "Big Data Science"
    Instructor: Dr. Anasse Bari
    Courant Institute of Mathematical Sciences, New York University
    Covers a variety of analytics algorithms and applications (dimensionality reduction, feature selection and
    extraction, similarity measures, decision trees, clustering, classification, recommender systems), as well
    as an introduction to large-scale file systems and MapReduce (Hadoop, Apache Mahout, Apache Spark)

Teaching Assistant   Winter 2014
    "Techniques in Computational Linguistics"
    Instructor: Prof. Dr. Alexander Koller
    Department of Linguistics, University of Potsdam
    Syllabus includes different parsing algorithms for context–free grammars; fundamental statistical methods
    such as N–gram language models, hidden Markov models, and probabilistic context–free grammars

Teaching Assistant   Summer 2013
    "Self–reflection and Planning"
    Department of Linguistics, University of Potsdam
    Designed and held a seminar intended to enhance first–year university experiences by imparting the
    basics of academic organizational structure and learning strategies (examined)

Languages

German	English	French
Native	Fluent	Working knowledge