ARTUR KULMIZEV

Engelska parken, Thunbergsv. 3 H, Box 635 ⋄ 751 26 Uppsala, Sweden

(+34) 666 71 64 07 ⋄ artur.kulmizev@lingfil.uu.se

EDUCATION

Uppsala University; Sweden   *2019-23*
PhD in Computational Linguistics
Project: multilingual morphosyntactic parsing
Supervisors: Joakim Nivre and Anders Søgaard

University of the Basque Country; Spain   *2017-18*
MSc in Language and Communication Technology
Year two of EM-LCT program.

University of Groningen; Netherlands   *2016-17*
MA in Language and Communication Technology
Year one of EM-LCT program.

University of Massachusetts-Amherst; USA   *2008-12*
BA in English Language and Literature
Cum Laude
Certificate in Technical Writing and Professional Communication
Minor in Spanish
Minor in Information Technology

HONORS AND SCHOLARSHIPS

GUF GRANT   *2017*
Awarded to outstanding Masters students at the University of Groningen.
EM-LCT Consortium Scholarship   *2016*
Awarded to motivated students enrolled in the EM-LCT program.
Dean's List   *2008-12*
Awarded to Bachelor's students who maintain at least a 3.5 GPA per semester.
John and Abigail Adams Scholarship   *2008-12*
Awarded to in-state Massachusetts students who maintain at least a 3.5 per year.

PUBLICATIONS

What can we learn from semantic tagging?   *2018*
Abdou, Kulmizev, Ravishankar, Abzianidze, Bos. Empirical Methods in Natural Language Processing
(EMNLP) 2018. - *forthcoming*
**AffecThor at SemEval-2018 Task One: A cross-linguistic approach
to sentiment intensity quantification in tweets**   *2018*
Abdou, Kulmizev, Ginés. Proceedings of the 12th International Workshop on Semantic Evaluation
(SemEval-2018).
Discriminator at SemEval-2018 Task Ten: Zero-shot Discrimination   *2018*
Kulmizev, et al. Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-
2018).
MGAD: Multilingual Generation of Analogy Datasets.   *2018*
Abdou, Kulmizev, Ravinshankar. Language Resources and Evalauation Conference (LREC) 2018.
The Power of Character N-Grams in Native Language Detection.   *2017*

Kulmizev, et al. 12th Workshop on Innovative Uses of NLP for Building Educational Applications. EMNLP 2017.

MASTER'S THESIS

Title: *Multilingual Word Embeddings and Their Utility in Cross-lingual Learning* (or lack thereof...)

- Proposed and investigated numerous supervised and unsupervised algorithms for extending bilingual word embedding mappings to a multilingual case, where a variable number of languages could exist in the same vector space.

- Evaluated the utility of word embedding alignment as it relates to cross-lingual part-of-speech tagging and dependency parsing. Conducted an extensive experiment to this end, incorporating numerous languages, sources of data, embedding alignments, and input representations.

RELEVANT ACADEMIC EXPERIENCE

Poster presentation at TABUDAG   *Groningen, NL — 2018*

Presented research regarding findings in multilingual transfer learning for dependency parsing enabled via multilingual word embeddings.

Poster Presentation at LREC 2018   *Miyazaki, JP — 2018*

· Presented work on generation of multilingual analogy datasets for word embedding evaluation at the 11th edition of Language Resources and Evaluation Conference.

SemEval-2018 Task 1: Affect in Tweets   *2018*
Winning group

· Implemented an ensemble of deep LSTMs and gradient-boosted classifiers using sentence, word, and character-level representations. Applied for sentiment affect regression and classification in Arabic, English, and Spanish Tweets.
· Won ordinal classification tasks for Arabic and Spanish.

SemEval-2018 Task 10: Capturing Discriminative Attributes   *2018*

· Explored the performance of various linear and non-linear classifiers in detecting discriminative attributes encoded in numerous sets of word embeddings, using techniques such as embedding retrofitting, distance thresholding, and sentence embedding.

2017 Shared Task on Native Language Identification   *2017*
Winning group

· Employed a character-level SVM (among other methods) in order to predict the native language of an author given a sample of their writing.
· Presented group-best results at EMNLP 2017 Poster Session in Copenhagen, Denmark.

ESSLLI   *Toulouse, FR — 2017*

· Attended the 29th European Summer School in Logic, Language, and Information.

Lot Summer School   *Leiden, NL — 2017*

· Attended the Landelijke Onderzoekschool Taalwetenschap.

WORK EXPERIENCE

University of Massachusetts, Amherst
Learning Management System Administrator

2012-2016
Amherst, Massachusetts, USA

· Front-end administration for Blackboard Learn and various building blocks.
· Management of institution-wide transition from self-hosted environment to managed hosting by Blackboard.
· Technical project management and logistical planning for online teaching and learning environments.
· Testing and implementation of service packs, cumulative patches, building blocks and integrations of other applications and services.

RELEVANT SKILLS

Programming Languages	Python, R, Java, Ruby
Machine Learning Toolkits	Tensorflow, PyTorch, Keras, scikit-learn, WEKA
Other	numpy; scipy; pandas; NLTK; gensim
	technical writing and documentation; git; LINUX/UNIX;
	web design (Bootstrap, Jekyll); data visualization
	(ggplot2, matplotlib, seaborn); LaTeX

LANGUAGES

English	Native-level
Russian	First language
Spanish	Written & Oral Proficiency
Dutch	Written & Oral Proficiency (B1)
Basque	Basic (A1)