

YURY ZEMLYANSKIY

+1 (650) 8420735 · yury.zemlyanskiy@usc.edu

University of Southern California, Los Angeles, CA 90089

RESEARCH INTERESTS

Machine Learning, Natural Language Processing, Representation Learning for Text and Entities.

EDUCATION

University of Southern California, Los Angeles, CA

2017-Present

PhD in Computer Science

Advisor: **Prof. Fei Sha**

Current GPA: 3.98/4.00

Saint Petersburg State University, St. Petersburg, Russia

2007-2012

Specialist Degree (5 years) in Computer Science

Advisor: **Prof. Alexander Vakhitov**

Thesis: “Variational approach for Super-resolution“

GPA: 3.87/4.00

RESEARCH EXPERIENCE

Google

March - October 2021

Student Researcher

Remote

- Hosts: William Cohen (wcohen@google.com), Nicholas FitzGerald (nfitz@google.com)
- MentionMemory: incorporate information from the entire corpus into Transformer to solve open-domain knowledge-intensive tasks.

Google

June - November 2020

Research Intern & Student Researcher

Remote

- Hosts: Joshua Ainslie (jainslie@google.com), Ilya Eckstein (ilyaeck@google.com)
- [ReadTwice](#) method: augment BERT model with memory to handle extremely large inputs.

Google

June - September 2019

Research Intern

Mountain View, CA

- Hosts: Ilya Eckstein (ilyaeck@google.com), Sudeep Gandhe (srgandhe@google.com)
- DOCENT, a family of models to learn entity representations from large document collections.
- [Reviews2Movielens](#) – a new multi-document, multi-label dataset created by joining Amazon Movie Reviews and MovieLens.

Facebook

February 2014 - May 2017

Research Scientist @ Machine Translation (Applied Machine Learning)

Menlo Park, CA, USA

- Neural Network Machine Translation in production. ([MIT Technology Review](#), [Wired](#) and [Slator](#)).
- Add [RNN](#) and [Seq2Seq](#) support to the Caffe2 framework.

PUBLICATIONS

YZ, Joshua Ainslie, Michiel de Jong, Philip Pham, Ilya Eckstein, and Fei Sha. “ReadTwice: Reading Very Large Documents with Memories.”, *NAACL 2021*.

YZ, Sudeep Gandhe, Ruining He, Bhargav Kanagal, Anirudh Ravula, Juraj Gottweis, Fei Sha and Ilya Eckstein, “DOCENT: Learning Self-Supervised Entity Representations from Large Document Collections”, *EACL 2021*.

Lukas Ruff, **YZ**, Robert Vandermeulen, Thomas Schnake and Marius Kloft, “Self-Attentive, Multi-Context One-Class Classification for Unsupervised Anomaly Detection on Text” *ACL 2019, Florence, Italy*.

YZ and Fei Sha, “Aiming to Know You Better Perhaps Makes Me a More Engaging Dialogue Partner”, *CoNLL 2018, Brussels, Belgium*.

Matthias Eck, **YZ**, Joy Zhang and Alex Waibel, “Extracting Translation Pairs from Social Network Content”, *IWSLT 2014, South Lake Tahoe, CA*.

CONFERENCE TALKS

“DOCENT: Learning Self-Supervised Entity Representations from Large Document Collections”, PhD Intern Research Conference (lightning talk), Google / remote, July. 2020.

“From Documents to Aspects to Answers: A Multi-label, Multi-document, Open-Vocabulary Classification Task”, *NDS NYAS* (poster), The New York Academy of Sciences, New York, Nov. 2019.

“Aiming to Know You Better Perhaps Makes Me a More Engaging Dialogue Partner”, *WeCNLP* (lightning talk, poster), Facebook HQ, Menlo Park, CA, Sep. 2018.

PROFESSIONAL ACTIVITIES

Workflow chair for [AAAI-2020](#), Reviewer for ACL (2018, 2021), NAACL (2021).

WORK EXPERIENCE

Facebook, Inc

October 2012 - January 2014

Software Engineer @ Site Integrity

London, UK

- Anti-spam department, Site Integrity team protecting people on Facebook from spam, scams, account compromises, and other forms of abuse.
- Developed several classifiers (SVM, Random Forests, statistical tests) to detect and neutralize malicious content. The classifiers are used on every URL which appears on Facebook.

Hewlett Packard Labs

May - August 2012

Research Intern

St. Petersburg, Russia

- Image segmentation and text detection algorithms for document scans.

OpenCV & Itseez

April - August 2012

Software Engineer Intern

St. Petersburg, Russia

- Participated in Google Summer of Code 2012 as an intern on a OpenCV project working on SimpleFlow optical flow estimation algorithm.

Facebook*Software Engineer Intern @ Ads Backend*

September - December 2011

Palo Alto, CA, USA

- Back-end team of Facebook Ads department.

Algorithmic Biology Lab*Research Intern*

July - August 2011

St. Petersburg, Russia

- Algorithmic Mass Spectrometry.

Yandex*Software Engineer Intern & Software Engineer*

December 2009 - July 2011

St. Petersburg, Russia

- Back-end team for Yandex.Maps.

TEACHING EXPERIENCE

University of Southern California

August 2021 - December 2021

- Teaching Assistant for [CS567: Machine Learning](#) by professor [Haipeng Luo](#).

University of Southern California

August 2019 - December 2019

- Teaching Assistant for [CS675: Convex and Combinatorial Optimization](#) by professor [Shaddin Dughmi](#).

Academic Gymnasium, St. Petersburg

September 2009 - May 2011

- Taught an algorithms and data structures course for programming competitions.

AWARDS

Third Prize at ACM ICPC, North-Eastern European Regional Contest, 2008 and 2009

Scholarship of the Government of the Russian Federation

TECHNICAL STRENGTHS

Programming Languages:

Python, C++, Java, Haskell

Scientific Computing Frameworks:

JAX, PyTorch, Tensorflow

scikit-learn, scipy, numpy, matplotlib

Other:

TeX