

Alexander Johansen

COPENHAGEN · DENMARK

☎ (+45) 5360-1352 | ✉ alexander@herhjemme.dk | 🏠 alrojo.github.io | 📷 alrojo | 📺 alexrosejo | 🐦 @alexrosejo

Research Interest

I am curious about the underlying latent structure of language. I investigate language through mathematical constructs, bioinformatics, and chemistry. My recent work is focused on building neural components that can extrapolate and transferring natural language processing methods into bioinformatics and cheminformatics. Proteins are the most structurally complex and sophisticated molecules in existence. This structure has been developed and fine-tuned over billions of years of evolutionary history. The language of proteins is structured with many unknown rules of behavior. This makes proteins an interesting vessel for learning latent features.

Education

Technical University of Denmark

M.Sc. MATHEMATICAL MODELING AND COMPUTATION, GPA: 11.44/12.00

- **Nanyang Technological University**, Singapore — *Semester Abroad Fall 2015*
- **Honors program**, Supervised by Professor Ole Winther

Kongens Lyngby, Denmark

Sept. 2014 - Dec. 2016

Copenhagen Business School

B.Sc. BUSINESS ADMINISTRATION AND INFORMATION SYSTEMS, GPA: 10.70/12.00

- **Lincoln University**, Canterbury, New Zealand — *Semester Abroad Fall 2013*

Frederiksberg, Denmark

Sept. 2011 - Jun. 2014

Experience

Technical University of Denmark

RESEARCH ASSISTANT

Inspired by my time at Salesforce Research, I started a student based research lab with research meetings every Tuesday and one-on-ones every Thursday. More than 30 students (M.Sc. and Ph.D.) have participated in the lab in 2019 resulting in several paper submissions. To find research projects I utilize the many professors at the university for ideation and to find interesting datasets. The majority of my students have previously done projects with me in 02456 Deep Learning.

Kongens Lyngby, Denmark

Jan 2019 - Present

Salesforce

DEEP LEARNING RESEARCH, INTERN

Under the supervision of Richard Socher (Salesforce Research) I researched in probability based decision making (+100k impression blogpost, ACL workshop paper); multi-task learning for NLP; and mixture-of-experts using distributed computing, PyTorch, and TorchText.

172 University Ave, Palo Alto, CA

94301, USA

Jan 2017 - Jan 2018

Teaching

Technical University of Denmark

DEEP LEARNING, 02456

In 2018 and 2019 I have been the head TA with significant course material contributions. Half of the course is project based and I supervise the NLP, Bioinformatics, and RL projects; the most popular amongst students for 2018 and 2019.

Head TA

Fall 2016, Fall 2018, Fall 2019

MASTER THESIS SUPERVISION

I co-supervise 11 Master Thesis projects (an M.Sc. thesis is 5 months full-time). The thesis' have investigated formal languages, Levenhstein transformer, multi task learning, exploration in VQA, bio-, and cheminformatics; resulting in multiple journal submissions.

Co-Supervisor

Spring 2019, Fall 2019

SPECIAL COURSES

I have supervised 10 special course projects (a special course is 4 months part-time). Similar to M.Sc. thesis, the special courses have investigated topics within deep learning methods and applications.

Co-Supervisor

Spring 2019, Fall 2019

INTRO REINFORCEMENT LEARNING

Exercises from Chap 1-13 in Sutton & Barto and Homework 1-2 from UC Berkeley's Deep RL course. 9 Students (M.Sc. and Ph.D.).

Course Responsible

Spring 2019

DEEP REINFORCEMENT LEARNING

Adjusted version of UC Berkeley's Deep RL course, co-supervised with Ass. Prof. T. Herlau. 10 students (M.Sc. and Ph.D.).

Course Co-Responsible

Jun. 2019

Community

Deep Learning Copenhagen

FOUNDER

MeetUp

Nov. 2018 - PRESENT

Inspired by Richard Socher's public poster exam in CS224N I convinced Professor Ole Winther to do the same for our 02456 Deep Learning course. With students posters, company sponsored first prize, drinks, and pizza. Given the positive feedback, me and a bioinformatics Ph.D. student started a research lab and kept hosting events for our students. This has resulted in seven events by now, +1.5k participants, and multiple company sponsorships. (Event page <https://www.meetup.com/Deep-Learning-DTU/> student research lab <https://alrojo.github.io>).

Open Source

GOOGLE TENSORFLOW

contrib.seq2seq: #4761, #4686, #4382

TensorFlow tutorial (2k stars): <https://github.com/alrojo/tensorflow-tutorial>

Academic Reviews

2020	AAAI , Association for the advancement of artificial intelligence	Program committee
2019	CoNLL , Computational Natural Language Learning	Reviewer
2018	CoNLL , Computational Natural Language Learning	Reviewer
2017	NIPS , Neural Information Processing Systems	Assisted review
2017	ICML , International Conference on Machine Learning	Assisted review

Journal Publications

An introduction to deep learning on biological sequence data: examples and solutions

PUBLISHED

BIOINFORMATICS (IF: 4.5) VOLUME 33, ISSUE 22, PAGES 3685-3690, OXFORD UNIVERSITY PRESS

<https://academic.oup.com/bioinformatics/article/33/22/3685/4092933>

2017

V. Jurtz, [A. Johansen](#), M. Nielsen, J. Armenteros, H. Nielsen, C. Sønderby, O. Winther and S. Sønderby

Conference Publications

Deep recurrent conditional random field for protein secondary structure prediction

ORAL

ACM CONFERENCE ON BIOINFORMATICS, COMPUTATIONAL BIOLOGY, AND HEALTH INFORMATICS

<http://delivery.acm.org/10.1145/3110000/3107489/p73-johansen.pdf>

2017

[A. Johansen](#), C. Sønderby, S. Sønderby and O. Winther

A deep learning approach to adherence detection for type 2 diabetics

POSTER

IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC)

<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7471776>

2017

A. Mohebbi, T. Aradóttir, [A. Johansen](#), H. Bengtsson, M. Fraccaro, M. Mørup

Epileptiform spike detection via convolutional neural networks

POSTER

IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING

<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8037462>

2016

[A. Johansen](#), J. Jin, T. Maszczyk, J. Dauwels, S. Cash and M. Westover

Workshop and Abstract Publications

Measuring arithmetic extrapolation performance

POSTER

NEURIPS WORKSHOP ON SCIENCE MEETS ENGINEERING OF DEEP LEARNING

<https://arxiv.org/abs/1910.01888>

2019

A. Madsen, [A. Johansen](#)

Language modeling for biological sequences — curated datasets and baselines

POSTER

NEURIPS WORKSHOP ON LEARNING MEANINGFUL REPRESENTATIONS OF LIFE

2019

J. Armenteros, [A. Johansen](#), O. Winther, H. Nielsen

Learning the language of life

ORAL

INTELLIGENT SYSTEMS FOR MOLECULAR BIOLOGY / EUROPEAN CONFERENCE ON COMPUTATIONAL BIOLOGY

https://orbit.dtu.dk/files/193584092/Learning_the_language_of_life_abstract.pdf

2019

J. Armenteros, [A. Johansen](#), O. Winther, H. Nielsen

Learning when to skim and when to read

ACL WORKSHOP ON REPRESENTATION LEARNING FOR NLP

[HTTPS://ARXIV.ORG/ABS/1712.05483](https://arxiv.org/abs/1712.05483)

[A. Johansen](#), R. Socher

POSTER

2017

Neural machine translation with characters and hierarchical encoding

NIPS RECURRENT NEURAL NETWORK SYMPOSIUM

[HTTPS://ARXIV.ORG/ABS/1610.06550](https://arxiv.org/abs/1610.06550)

[A. Johansen](#), J. Hansen, E. Obeid, C. Sønderby and O. Winther

POSTER

2016

Current Projects

Neural arithmetic units

SUBMITTED TO INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS

[HTTPS://OPENREVIEW.NET/FORUM?ID=H1gN0eHKPS](https://openreview.net/forum?id=H1gN0eHKPS)

A. Madsen, [A. Johansen](#)

UNDER REVIEW, Ratings: 8/3/6/6

Autoencoding undirected molecular graphs with neural networks

SUBMITTED TO JOURNAL OF CHEMICAL INFORMATION AND MODELING (IF: 3.9)

[HTTPS://ALROJO.GITHUB.IO/MEDIA/PUBLICATIONS/MOLECULES/PREPRINT.PDF](https://alrojo.github.io/media/publications/molecules/preprint.pdf)

J. Olsen, P. Christensen, M. Hansen, [A. Johansen](#)

UNDER REVIEW

Prediction of GPI-Anchored proteins with pointer neural networks

SUBMITTED TO PROTEINS: STRUCTURE, FUNCTION, AND BIOINFORMATICS (IF: 2.5)

[HTTPS://WWW.BIORXIV.ORG/CONTENT/10.1101/838680V1.ABSTRACT](https://www.biorxiv.org/content/10.1101/838680v1.abstract)

M. Gislason, H. Nielsen, J. Armenteros*, [A. Johansen](#)* (*equal contribution)

UNDER REVIEW

Language modeling for biological sequences — curated datasets and baselines

TO BE SUBMITTED FOR BIOINFORMATICS (IF: 4.5) BY MEDIO JAN 2020

[HTTPS://GITHUB.COM/ALROJO/UNILANGUAGE/BLOB/MASTER/PREPRINT.PDF](https://github.com/alrojo/unilanguage/blob/master/preprint.pdf)

J. Armenteros*, [A. Johansen](#)*, O. Winther, H. Nielsen (*equal contribution)

TO BE SUBMITTED

Predicting recombinant gene expression with deep learning techniques

TO BE SUBMITTED FOR JOURNAL OF BIOTECHNOLOGY (IF: 3.1) BY MEDIO JAN 2020

H. Martiny, J. Armenteros, [A. Johansen](#), J. Salomon, H. Nielsen

TO BE SUBMITTED

Short term blood glucose prediction based on continuous glucose monitoring data

TO BE SUBMITTED FOR IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC) BY 1ST JAN 2020

A. Mohebbi, [A. Johansen](#), N. Hansen, P. Christensen, M. Jensen, J. Tarp, H. Bengtsson, M. Mørup

TO BE SUBMITTED

Patents

Probability-Based Guider

US PATENT APP. 15/853,530

[A. Johansen](#), B. McCann, J. Bradbury, R. Socher

PENDING

2017

Deep Neural Network-Based Decision Network

US PATENT APP. 15/853,570

[A. Johansen](#), B. McCann, J. Bradbury, R. Socher

PENDING

2017

Technical Skills

Programming Python, Matlab, SQL, Java

ML Frameworks PyTorch, TensorFlow, Theano, Lasagne, CUDA

Others Linux, Docker, Vim, IPythonNotebook, Google Colab, Git, Github, AWS S3, AWS EC2, \LaTeX