

# ALEKSANDAR DIMITRIEV

## CURRICULUM VITAE

 [alekdimi](#)       [dimitriev.aleksandar@gmail.com](mailto:dimitriev.aleksandar@gmail.com)

(+386)40665046      Aljaževa ulica 6, 1000 Ljubljana, Slovenia

 [alekdimi](#)       [thenaivebayesian.com](http://thenaivebayesian.com)

## RESEARCH INTERESTS

Machine Learning, Bayesian Statistics, Data Mining, Computer Vision, Bioinformatics, Social Network Analysis.

## EDUCATION

July 2016 – Sept. 2016      Stanford University, United States

*Visiting  
Researcher*

ASEF Fellowship Awardee: Worked on extracting large-scale protein-protein interaction networks from publicly available PubMed paper abstracts using natural language processing and predictive models.  
Advisor: Prof. Jure LESKOVEC

2014-      University of Ljubljana, Slovenia

*MSc in Computer  
Science*

GPA: 9.9/10 · Dean's and Rector's Commendations for outstanding academic success  
Faculty of Computer and Information Science  
Advisor: Prof. Erik ŠTRUMBELJ

2011-2014      University of Ljubljana, Slovenia

*BSc in Computer  
Science*

GPA: First year: 7.7/10 · Second year: 8.6/10 · Third year: 9.7/10 · Thesis: 10  
Faculty of Computer and Information Science  
Thesis: *A Markov Random Field Based Autonomous Image Segmentation*  
Advisor: Prof. Matej KRISTAN

2007-2011      Orce Nikolov, Macedonia

*High School*

GPA: 5.0/5 · Class Valedictorian  
Natural Sciences Track, Bilingual Program (English and Macedonian)

## PUBLICATIONS

*In Preparation*      **Large Scale Extraction of Protein Protein Interactions Using Data Programming**

We analyzed 26 million publicly available abstracts on PubMed for protein protein interactions. We trained a supervised classifier to predict whether a candidate interaction is a PPI given all the sentences in which it occurs. To obtain ground truth for the classifier, we used data programming, and wrote labeling functions. We evaluated our predictions on two well known protein interaction databases, which show promise for data programming for large unlabeled data sets.

Authors: Aleksandar DIMITRIEV, Stephen BACH, Rok SOSIĆ, Jure LESKOVEC

Dec. 2016      **Approximate Bayesian Binary, Ordinal Regression with Structured Uncertainty in the Inputs**

*NIPS '29  
(Advances in  
Neural Inf. Proc.  
Systems)  
Workshop on  
Approximate  
Inference, 2016*

We tackled the problem of non-iid noisy data by marginalizing over all possible data sets in both training and testing. To achieve this, we used Monte Carlo sampling of training and test data sets, assuming the structure of the uncertainty is known, with comparable speed to ordinary models. In sports data, where considerable noise in the beginning of the season is present, our model significantly outperformed its noiseless counterparts, especially at the start of a new season.

Authors: Aleksandar DIMITRIEV, Erik ŠTRUMBELJ

*Feb. 2016*      **A regularization-based approach for unsupervised image segmentation**

Improved and extended version of the BSc thesis. We designed and evaluated a novel image segmentation algorithm that requires no user input or information about the number of segments, and is comparable to the state of the art. Our approach works on the superpixel level, by first oversegmenting the regions, training an SVM for each, and then merging them based on their similarity. To take into account the structural information that is present in an image, e.g. similarity of neighboring regions, a Markov random field is superimposed on the superpixels.

Authors: Aleksandar DIMITRIEV, Matej KRISTAN

*Journal of  
Electrical  
Engineering and  
Computer Science*

*Oct. 2015*      **Learning from microarray gene expression data**

We analyzed several feature selection techniques and machine learning classifiers on four gene microarray data sets to evaluate their ability to cope with very sparse high-dimensional data. We ascertained that choosing which feature selection technique to use is less important than choosing the classifier. In addition, models trained on one dataset were successfully predicting on other datasets, provided the genes, i.e. the feature space, is the same, which indicates that the models generalized well.

Authors: Aleksandar DIMITRIEV, Zoran BOSNIĆ

*18<sup>th</sup> Information  
Society  
Multiconference,  
2015*

*Sept. 2014*      **The Visual Object Tracking VOT2014 challenge results**

Submitted a novel short-term visual tracker that was competitive with the state of the art on the VOT toolkit, which tests the each tracker on a variety of challenging videos.

Authors: Matej KRISTAN, ... , Aleksandar DIMITRIEV, et al.

*ECCV '13  
(European  
Conference on  
Computer Vision)  
Workshop paper,  
2014*

*Sept. 2014*      **Advanced computer vision methods for autonomous navigation of a robotic marine vessel**

An extended abstract that highlighted the key achievements of a project for improving the autonomy of a sea robot, among which was an early prototype of the aforementioned segmentation algorithm.

Authors: Aleksandar DIMITRIEV, Franci OVEN, Tina STRGAR, Aleš CERNE, Jošt NOVAK, Duško VRANAC, Matej KRISTAN, Janez PERŠ

*23<sup>rd</sup>  
Electrotechnical  
and Computer  
Science Conference,  
2014*

**TEACHING***Sept. 2015 –*      **Social and Information Network Analysis**

Elective course for undergraduate and graduate students that is held at the University of Ljubljana and precisely follows Prof. **Jure Leskovec**'s Stanford version. Instructed students by mentoring and guiding course projects, summarizing lectures, leading discussions, maintaining the course website, and grading students' homework and project reports.

*Teaching  
Assistant*

**PROFESSIONAL EXPERIENCE***Feb. 2016 –*      **SALVIOL GLOBAL ANALYTICS** — Ljubljana, Slovenia

Part-time internship to develop a framework in *R* (and integrate in *Java*) for fraud detection in insurance data in the property and automobile industry. Using machine learning, we modeled the probability of fraud during insurance policy creation, long before a claim for damages is filed, to prevent fraud right at the beginning.

*Data Scientist  
(R, Java)*

*April 2015 –  
June 2015*      **BIOINFORMATICS LABORATORY** — UNIVERSITY OF LJUBLJANA, SLOVENIA

Developed a bioinformatics add-on for data fusion using matrix factorization in **Orange** for the purpose of showcasing **Orange** at the University of Toronto.

*Software  
development  
(Python)*

*Dec. 2014 –  
July 2015*      **INSTITUTE "JOŽEF STEFAN"** — Ljubljana, Slovenia

Maintained and improved **Elycite**, a large-scale data mining tool for text corpora built on **QMiner**, which uses (semi-)supervised learning for predictive modelling on sentences.

*Internship  
(Javascript, C++)*

April 2014 – **HARPHA SEA** — Koper, Slovenia  
 Sept. 2014

Industry  
 collaboration  
 (MATLAB)

Worked on a joint **research project** with the Faculty of Computer Science, University of Ljubljana, on developing a segmentation algorithm, among other techniques, to aid a marine robotic vessel in collision detection and robustness.

#### PROGRAMMING LANGUAGES

Expert	PYTHON, R, MATLAB
Experienced	JAVA, C, JAVASCRIPT
Familiar	SQL, L <sup>A</sup> T <sub>E</sub> X, WEB PROGRAMMING

#### OTHER INFORMATION

Selected Awards  
 & Honors

2016 · RECTOR'S COMMENDATION — Annual award for outstanding academic success, given by the President of the University of Ljubljana to 10 students out of 40,000 enrolled, Ljubljana, Slovenia.

2016 · **ASEF FELLOWSHIP** — Research fellowship in Computer Science for a 10-week visit at Stanford University under the supervision of Prof. Leskovec, Stanford, CA.

2016 · SILVER MEDAL — 1<sup>st</sup> "Kangaroo" University of Ljubljana Mathematics Competition, Ljubljana, Slovenia.

2015 · DEAN'S COMMENDATION — For exemplary academic success in the 2014–2015 school year, Faculty of Computer and Information Science, University of Ljubljana, Slovenia.

2015 · 3<sup>rd</sup> PLACE — **DREAM Olfaction Prediction Challenge**, Team *Biolab Ljubljana*.

2011 · VALEDICTORIAN — "Orce Nikolov" High School, Class of 2011, Skopje, Macedonia.

2011 · 2<sup>nd</sup> PLACE — 2<sup>nd</sup> State Competition in Astronomy, Skopje, Macedonia.

2009 · BRONZE MEDAL — 11<sup>th</sup> Macedonian Mathematical Olympiad, Skopje, Macedonia.

2007 · 3<sup>rd</sup> PLACE — 31<sup>st</sup> State Competition in Physics, Skopje, Macedonia.

2005 · 1<sup>st</sup> PRIZE — 42<sup>nd</sup> Violin State Competition, Skopje, Macedonia.

Official Scores

Nov. 2015 · GRE — Quantitative Reasoning: 170/170, Verbal Reasoning: 168/170, Analytical Writing 4.0.  
 Nov. 2015 · TOEFL iBT — 119/120.

Languages

MACEDONIAN (native) · ENGLISH (fluent) · SLOVENIAN (fluent) · GERMAN (limited proficiency)

Selected  
 Certificates

2015 · STATISTICAL LEARNING — Stanford University, *Stanford Online*.

2015 · MATRIX ALGEBRA AND LINEAR MODELS — Harvard University, *edX*.

2015 · DATA SCIENCE AND MACHINE LEARNING ESSENTIALS — Microsoft, *edX*.

2015 · SCALABLE MACHINE LEARNING — UC Berkeley, *edX*.

2015 · INTRODUCTION TO BIG DATA WITH APACHE SPARK — UC Berkeley, *edX*.

2015 · INTRODUCTION TO COMPUTATIONAL THINKING AND DATA SCIENCE — MIT, *edX*.

2014 · DATA ANALYSIS AND STATISTICAL INFERENCE — Duke University, *Coursera*.

2014 · DATA ANALYSIS — Johns Hopkins University, *Coursera*.

2014 · COMPUTING FOR DATA ANALYSIS — Johns Hopkins University, *Coursera*.

2014 · MACHINE LEARNING — Stanford University, *Coursera*.