

ALEKSANDAR DIMITRIEV

CURRICULUM VITAE

 [alekdimi](#)  alekdimi@utexas.edu
 [alekdimi](#)  thenaivebayesian.com

RESEARCH INTERESTS

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

EDUCATION

2017- **University of Texas at Austin, United States**

*PhD
Risk Analysis &
Decision Making*

Graduate School Recruitment Fellowship & McCombs Dean's Fellowship holder
McCombs School of Business · Department of Information, Risk, and Operations Management (IROM)
Statistics Track.

July-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

2014-2017 **University of Ljubljana, Slovenia**

*MSc
Computer Science*

GPA: 9.9/10 · Dean's and Rector's Commendations for outstanding academic success
Faculty of Computer and Information Science
Thesis: *Modelling Multivariate Discrete Data With Latent Gaussian Processes*
Advisor: Prof. Erik ŠTRUMBELJ

2011-2014 **University of Ljubljana, Slovenia**

*BSc
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

Feb. 2017 **Predicting human olfactory perception from chemical features of odor molecules**

*Science
Volume 355
Issue 6327,
pages 820-826*

Report of the *DREAM Olfaction Prediction Challenge*, which was a crowd-sourced effort to predict whether a given molecule will have a perceived odor and what odor it will produce among 19 semantic descriptors (sweet, musky, garlic, ...) as well as odor intensity and pleasantness. Our team *Biolab Ljubljana* achieved 3rd place using an ensemble of machine learning methods, including random forest, gradient-boosted trees, and ridge regression.

Authors: Andreas KELLER, ... , DREAM Olfaction Prediction Consortium (Agnieszka Kitlas GOLISKA, Aleksandar DIMITRIEV, ...), et al.

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**

Journal of
Electrical
Engineering and
Computer Science

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

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

18th Information
Society
Multiconference,
2015

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

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

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

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.

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

23rd
Electrotechnical
and Computer
Science Conference,
2014

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

TEACHING

Spring 2018 **Elementary Business Statistics (STA 309) — Austin, TX**

Teaching
Assistant

Core statistics course for undergraduate students at McCombs School of Business, taught by Prof. **Betsy Greenberg** and **Elizabeth Moliski**.

Fall 2017 **Statistics and Modeling (STA 371G) — Austin, TX**

Teaching
Assistant

Core statistics course for undergraduate students at McCombs School of Business, taught by Prof. **Paul Damien**.

Fall 2015 &
Fall 2016 **Social and Information Network Analysis — Ljubljana, Slovenia**

Teaching
Assistant

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.

PROFESSIONAL EXPERIENCE

- March 2016 – June 2017 **ZEMANTA** — Ljubljana, Slovenia
- Data Scientist*
(Python, Go, Hadoop)
- Full-time internship to develop a fraud detection framework. The goal was detecting non-intentional traffic on websites or apps that serve real-time native advertisements by Zemanta. Among other things, we used HADOOP in Go to automatically process large-scale data on a weekly basis and construct a co-visitation network of shared IPv4 traffic between sites that can be clustered to generate suspicious sites.
- Feb. 2016 – Feb. 2017 **SALVIOL GLOBAL ANALYTICS** — Ljubljana, Slovenia
- Data Scientist*
(R, Java)
- Part-time internship to develop a framework in R and integrate it 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.
- April 2015 – June 2015 **BIOINFORMATICS LABORATORY** — UNIVERSITY OF LJUBLJANA, SLOVENIA
- Software Developer*
(Python)
- Developed a bioinformatics add-on for data fusion using matrix factorization in Orange for the purpose of showcasing Orange at the University of Toronto.
- Dec. 2014 – July 2015 **INSTITUTE "JOŽEF STEFAN"** — Ljubljana, Slovenia
- Intern*
(Javascript, C++)
- Maintained and extended Elycite, a large-scale data mining tool for text corpora built on QMiner, which uses (semi-)supervised learning for predictive modelling on sentences.
- April 2014 – Sept. 2014 **HARPHA SEA** — Koper, Slovenia
- 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

<i>Interpreted</i>	PYTHON, R, MATLAB, JAVASCRIPT
<i>Combined</i>	JAVA, C, GO
<i>Other</i>	HADOOP, SQL, L ^A T _E X

OTHER INFORMATION

- Selected Awards & Honors*
- 2017- · MCCOMBS DEAN'S FELLOWSHIP — University of Texas at Austin.
- 2017- · GRADUATE SCHOOL RECRUITMENT FELLOWSHIP — University of Texas at Austin.
- 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 — 1st "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 · 3rd PLACE — DREAM Olfaction Prediction Challenge, Team Biolab Ljubljana.
- 2011 · VALEDICTORIAN — "Orce Nikolov" High School, Class of 2011, Skopje, Macedonia.
- 2011 · 2nd PLACE — 2nd State Competition in Astronomy, Skopje, Macedonia.

- 2009 · BRONZE MEDAL — 11th Macedonian Mathematical Olympiad , Skopje, Macedonia.
 2007 · 3rd PLACE — 31st State Competition in Physics, Skopje, Macedonia.
 2005 · 1st PRIZE — 42nd 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
 FRENCH · Intermediate
 GERMAN · Limited
 SERBO-CROATIAN · Working 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*.

February 2, 2018