

DIMITRIS SPATHIS

FN01, William Gates Building
Dept. of Comp. Science and Tech.
University of Cambridge
Cambridge, CB3 0FD, UK
ds806@cam.ac.uk
www.cl.cam.ac.uk/~ds806

EDUCATION

- 2017- PhD in Computer Science
University of Cambridge, UK
Advisor: Prof. Cecilia Mascolo
Topic: Representation learning to model human behavior and health with high-dimensional mobile sensing data
- 2015-17 MSc in Computational Intelligence
Aristotle University of Thessaloniki, Greece
Advisor: Prof. Anastasios Tefas
Thesis: Interactive dimensionality reduction and manifold learning
- 2011-15 BSc in Computer Science
Ionian University, Greece
Advisor: Prof. Panayiotis Vlamos
Thesis: Diagnosing respiratory diseases with machine learning

EXPERIENCE

- 2019 Ocado, Barcelona, Spain (Jun-Sep)
Data Scientist in Predictive Maintenance, Intern
Developed deep learning anomaly detection models for warehouse robot collision prevention, internal systems monitoring, and industrial conveyor belts
- 2018- University of Cambridge, UK
Teaching Assistant
Supervised bachelor theses & tutored students for the courses *Scientific Computing, Machine Learning & Real-World Data*, and *Mobile & Sensor Systems*
- 2017 Qustodio, Barcelona, Spain (Jun-Sep)
Data Scientist in Product, Intern
Built data pipelines to automate C-level KPI reporting, conducted A/B tests, custom cohort analyses and clustering to model customer retention and churn
- 2016 Telefonica Research, Barcelona, Spain (Sep-Dec)
Research Intern, hosted by Dr. Ilias Leontiadis
Along with a research team, proposed a new deep learning method to categorize unseen text and presented it at a workshop of ACL 2017 conference in Vancouver
- 2015 Aristotle University of Thessaloniki, Greece (Oct-Feb)
Teaching Assistant, Prof. Anastasios Tefas
Provided assistance to the undergraduate course Numerical Analysis
- Center for Research and Technology Hellas (CERTH), Greece (Jul-Sep)
Research Engineer Intern

Designed and developed the prototype mobile app for the EU Horizon asthma modelling research project *myAircoach*

2013-15

Social Informatics Lab, Ionian University, Greece

Undergraduate Researcher, Prof. Katia Kermanidis & Spyros Sioutas

Published work in international conferences and journals, on topics ranging from interactive geographic information systems to computational linguistics

DISTINCTIONS

2019

KDD 2019 Student Travel Award, ACM SIGKDD
\$1000 grant to attend *KDD'19 (USA)*

Graduate Research Travel Grant, Jesus College, University of Cambridge
£500 grant to attend *PervasiveHealth'19 (Italy)*

2018

PhD Open Day, Facebook office, London

Selected to attend internal research presentations along with 60 PhD students from the UK and presented one of the 10 shortlisted posters

AI Summer School, Microsoft Research, Cambridge

Selected to attend an intensive one-week program along with 100 AI PhD students from Europe

2017

Embricos Trust Scholarship, Jesus College, University of Cambridge & EPSRC Doctoral Training Partnership (DTP) Grant, University of Cambridge
Full PhD scholarship

2016

Marie Skłodowska-Curie Research and Innovation Staff Exchange (RISE), EU Horizon2020 Project "*ENCASE*": online fraud & harassment detection

2015

Admission Scholarship to MSc, Aristotle University of Thessaloniki
Full tuition fees coverage for the 1st semester

Investigative Data Journalism Contest, *Innovathens*

Glocal News web-app, 3rd prize

2014

Erasmus+ multinational project management training in Iceland

Full financial support by EU *YouthInAction*, *Youthnet Hellas*

2013

Educational Trip to Silicon Valley. Selected among 250+ students nationwide.

Attended courses at *Stanford University* and visited *UC Berkeley*, *Google*, *Facebook*, *Twitter* and *Coursera* offices. Read more: <http://medium.com/p/37e8d2ca0421>

PUBLICATIONS

— Peer reviewed journals, conferences or workshops

1. Brown*, C., Chauhan*, J., Grammenos*, A., Han*, J., Hasthanasombat*, A., Spathis*, D., Xia*, T., Cicuta, P., & Mascolo, C. (2020). Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data. *International Conference on Knowledge Discovery and Data Mining (KDD), Virtual Event. (Oral presentation)* <http://doi.org/d683>

*equal contribution, listed in alphabetical order

2. Spathis D., Rodriguez, S., Farrahi, K., Mascolo, C., & Rentfrow, J. (2019). Sequence Multi-task Learning to Forecast Mental Wellbeing from Sparse Self-reported Data. *International Conference on Knowledge Discovery and Data Mining (KDD), Anchorage, USA. (Oral presentation, top 6% of submissions)* <http://doi.org/gf7nbh>
3. Spathis, D., Rodriguez, S., Farrahi, K., Mascolo, C., & Rentfrow, J. (2019). Passive mobile sensing and psychological traits for large scale mood prediction. *International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), Trento, Italy.* <http://doi.org/c7hk>
4. Spathis, D., Passalis, N., & Tefas, A. (2019). Interactive dimensionality reduction using similarity projections. *Knowledge-Based Systems, 165:77-91.* <http://doi.org/cxbm>
5. Spathis, D., Passalis, N., & Tefas, A. (2018). Fast, Visual and Interactive Semi-supervised Dimensionality Reduction. *European Conference on Computer Vision (ECCV), Workshop on Compact and Efficient Feature Representation and Learning in Computer Vision, Munich, Germany.* <http://doi.org/cz6d>
6. Spathis, D., Vlamos, P. (2017). Diagnosing Asthma and Chronic Obstructive Pulmonary Disease with Machine Learning. *Health Informatics Journal, 25(3): 811:827.* Epub 2017, issue published in 2019. <http://doi.org/cbzh>
7. Serrà, J., Leontiadis, I., Spathis, D., Stringhini, G., Blackburn, J., Vakali, A. (2017). Class-based Prediction Errors to Detect Hate Speech with Out-of-vocabulary Words. *Annual Meeting of the Association for Computational Linguistics (ACL), Workshop on Abusive Language Online, Vancouver, Canada.* <http://doi.org/b94p>
8. Charalampakis, B., Spathis, D., Kouslis, E., & Kermanidis, K. (2016). A comparison between semi-supervised and supervised text mining techniques on detecting irony in greek political tweets. *Engineering Applications of Artificial Intelligence, 51:50–57.* <http://doi.org/bzxq>
9. Charalampakis, B., Spathis, D., Kouslis, E., & Kermanidis, K. (2015). Detecting Irony on Greek Political Tweets: A Text Mining Approach. *International Conference on Engineering Applications of Neural Networks (EANN), 17:1–5.* <http://doi.org/bzxr>
10. Spathis, D., Mouratidis, T., Sioutas, S., & Tsakalidis, A. (2014). Glocal News: An Attempt to Visualize the Discovery of Localized Top Local News, Globally. *International Conference on Conceptual Modeling, Workshop on Legal and Social Aspects in Web Modeling, Hong Kong, China.* Springer LNCS 8697. <http://doi.org/bzxs>

PROJECTS

- Open source code and contributions on Github <https://github.com/sdimi>
- 1. Covid-19 Sounds App. Crowdsourcing coughs to detect the coronavirus with ML. (60K+ visits) <http://covid-19-sounds.org>
Featured in *BBC, Forbes, The Guardian, The Times, Slate, El País*, and more.
- 2. Map out your musical taste on Spotify. Published in *Cuepoint Magazine*. (25K+ visits) <https://medium.com/p/fe50c94b8af3>
“Nifty music-oriented dataviz” – (tweet) Paul Lamere, Director of Dev.Plat., Spotify

3. How do popular book authors use language differently?
Text mining *Game of Thrones*, *Harry Potter*, *Hunger Games* and *Lord of the Rings* books. (10K+ visits) <http://medium.com/p/100290c94242>
Tweeted by @Medium to 2 Million followers, featured in *Editor Picks*
4. Visualizing pop-culture TV references
Every pop-culture reference on the TV series *Community*, per episode, per character. (15K+ visits) <http://communitypoprefs.com>
"This is very cool" – (tweet) Josef Adalian, Editor, *New York Magazine*
5. Google News mashup on top of Google Maps (now defunct due to API closure)
<https://github.com/sdimi/glocalnews.js/>
Contacted by the U.S. Department of Health in order to use part of this application for their policy research on health news coverage
6. Share your kid's pictures online without worrying (*HackZurich 2018* project)
<https://devpost.com/software/patronus-k61iv4>
Mobile app with face recognition, age estimation, & emotion recognition to blur underage people in pictures or replace their face with emotion-based emoji

ACADEMIC SERVICE

- Member of the Program Committee (PC), Research Track
Intl. Conference on Knowledge Discovery and Data Mining (KDD 2020), ACM
- Paper reviewer for international journals and conferences

General *Scientific Reports*, *Nature*

ML/Data *Intl. Conference on Machine Learning (ICML)*, IMLS
Expert Systems with Applications, Elsevier
Intl. Conference on Knowledge Discovery and Data Mining (KDD), ACM
Neurocomputing, Elsevier
International Joint Conference on Artificial Intelligence (IJCAI)
Intl. Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE
Intl. Conference on World Wide Web (WWW), IW3C2/ACM
Engineering Applications of Artificial Intelligence, Elsevier
Intl. Conference on Web and Social Media (ICWSM), AAAI
Intl. Conference on Pattern Recognition (ICPR), IAPR/IEEE

HCI *Conference on Human Factors in Computing Systems (CHI)*, ACM
Proc. Interactive, Mobile, Wearable and Ubiquitous Technologies (Ubicomp), ACM

Misc. *Health Informatics Journal*, Sage
The Visual Computer, Springer
Computer Networks, Elsevier

SKILLS

- Everyday use over the past 5-6 years
Data Science — Python: Numpy, Pandas, Scipy, Jupyter, Statsmodels
Machine Learning/Deep Learning — Sklearn, Tensorflow, Keras, PyMC
Data Processing — Unix/Bash, Slurm/GPU clusters, SQL

Visualization — Matplotlib, Seaborn

Project Management/Collaboration Tools — Jira, Git, Slack, Overleaf, GDocs

— Confident or used in past projects

C++, Java/Android, PHP, Javascript, R, D3.js, Plotly, Node.js, Google Cloud ML, Spark, NetworkX, Igraph, Gephi, Matlab, NLTK, Gensim

— Languages

English (C2), Spanish (B2), German (B2), Greek (native)