

Georgios Spithourakis

Address: 1st Floor, 90 High Holborn,
WC1V 6LJ, London, United Kingdom

Email: geospith@gmail.com
Homepage: <http://geospith.github.io>

RESEARCH INTERESTS

- Areas: natural language processing, machine learning, computational linguistics
- Interests: modelling text/categorical/continuous data, conditional and grounded language modelling, language generation, conversational agents, neural networks, numerical reasoning, structured prediction, forecasting

RESEARCH EXPERIENCE

University College London, UK

PhD Candidate in Natural Language Processing, September 2013 – present (submitting in July 2018)

Neural networks for modelling continuous/categorical/text data • Number-aware language models for error detection, error correction and text prediction for clinical and scientific data • Organised public engagement

Amazon, Cambridge, UK

Research Scientist Intern, July 2016 – February 2017

Question answering with neural networks for dialogue systems (related to Alexa product)

Microsoft Research, Redmond, Washington, US

Research Intern, June 2015 – August 2015

Conversation modelling with neural networks (related to Cortana product)

Forecasting and Strategy Unit, National Technical University of Athens, Greece

Research Associate, September 2011 – August 2012

Supply chain forecasting • Deployed and evaluated forecasting simulation platform for teaching forecasting at MSc level • Drafted research proposals

Research Software Engineer, September 2009 – August 2011

Developed web-based tools for time-series forecasting

EDUCATION

University College London, UK

PhD Candidate in Natural Language Processing, submitting in July 2018

- Thesis: "Numeracy of Language Models: Joint Modelling of Words and Numbers"
- Supervised by Sebastian Riedel and Steffen Petersen
- Funded by the Farr Institute of Health Informatics Research Studentship

University College London, UK

MSc in Computational Statistics and Machine Learning, 2013

- Themes: Supervised, unsupervised & reinforcement learning • Information retrieval • Machine vision
- Dissertation: "Column generation in loopy models"
- Supervised by Sebastian Riedel
- Funded by the Foundation for Education and European Culture (IPEP) Scholarship

National Technical University of Athens, Greece

Diploma (BSc/MSc) in Electrical & Computer Engineering, 2011

- GPA: 8.58/10.0 (28th out of 325 in year of graduation); completed the 5-year degree in 4 years
- Themes: Control theory & signal processing (speech, image, video) • Computer Networks (web application development, protocols, topology, hardware) • Management (decision support systems, forecasting)
- Dissertation: "Application of aggregate-disaggregate forecasting methodology to non-intermittent demand data"
- Supervised by Vasilis Assimakopoulos
- Funded by the Estate of Panagiotis Triadafyllidis Scholarship

PROGRAMMING SKILLS

Proficient in Python (including Tensorflow)

Previous experience with:

- Programming languages: Java, Scala, VB.NET, C++, MATLAB, R
- Web & Mobile Application Development: HTML, CSS, JavaScript, ASP.NET, AJAX, XML, JSON, Android
- Apache: Solr/Lucene, Hadoop, Spark
- Databases: SQL, MySQL

TEACHING EXPERIENCE

Department of Computer Science, University College London, UK

Teaching Assistant

- Natural Language Processing (MSc module, directed by S. Riedel), autumn terms 2013 – 2018
- Object-Oriented Programming (BSc module, directed by D. Mohamedally), spring terms 2014 – 2015
- Mobile App Design (MSc module, directed by D. Mohamedally), autumn terms 2013 – 2015
- Introductory Programming (MSc module, directed by Dean Mohamedally), autumn terms 2013 – 2014

Senior Teaching Assistant

- Introductory Programming (MSc module, directed by D. Mohamedally), autumn term 2015

National Institute of Informatics (NII), Tokyo, Japan

Visiting Teaching Assistant/Project Advisor (in collaboration with UCL)

- Project “Platform for E-Health Analytics and Collaborative Studies using Hadoop”, (1 week) 2016

School of Electrical and Computer Engineering, National Technical University of Athens, Greece

Teaching Assistant

- Forecasting Techniques (MSc module, directed by V. Assimakopoulos), spring terms 2010 – 2012

MENTORING & STUDENT SUPERVISION

Alexander Hoyle. Detecting statements that can be fact checked. Dissertation for MSc in Machine Learning, UCL, co-supervised with Jeff Mitchel, to finish in 2018.

Finneas Jacob Robson Catling. Towards automated clinical coding. Dissertation for MSc in Business Analytics (specialisation in Computer Science), UCL, co-supervised with Sebastian Riedel, 2017.

Jiayu Sherri Lee. Question-to-statement conversion using pivot translations. Dissertation for MSc in Computer Science, UCL, co-supervised with Sebastian Riedel, 2017.

Benjamin Riedel. Simple yet powerful: A system for news article stance detection. Dissertation for MSc in Machine Learning, UCL, co-supervised with Sebastian Riedel and Isabelle Augenstein, 2017.

Dhruv Ghulati. Distant Supervision and Cost Sensitive Classification for Weakly Supervised Claim Detection. Dissertation for MSc in Computer Science, UCL, co-supervised with Sebastian Riedel and Isabelle Augenstein, 2016.

Emanuel Mountalas. Empirical application, evaluation and comparison of forecasting methods for the electrical system’s System Marginal Price: a case study in Greek electricity market. Dissertation for diploma (BSc/MSc) in Electrical Engineering, 2012.

Triantafyllos Nikolaos Theodorou. Investment portfolio management with trend forecasting methods. Dissertation for diploma (BSc/MSc) in Electrical Engineering, NTUA, 2012.

Andreas Desalermos. Application and evaluation of optimisation techniques for exponential smoothing parameters. Dissertation for diploma (BSc/MSc) in Electrical Engineering, NTUA, co-supervised with Fotios Petropoulos, 2011.

PUBLICATIONS

Conference Proceedings

G. Spithourakis and S. Riedel. Numeracy for Language Models: Evaluating and Improving their Ability to Predict Numbers. *ACL 2018* (under review, decision pending).

N. Mostafazadeh, C. Brockett, B. Dolan, M. Galley, J. Gao, **G. Spithourakis** and L. Vanderwende. Image-Grounded Conversations: Multimodal Context for Natural Question and Response Generation. *IJCNLP 2017*.

G. Spithourakis, I. Augenstein and S. Riedel. Numerically Grounded Language Models for Semantic Error Correction. *EMNLP 2016*.

J. Li, M. Galley, C. Brockett, **G. Spithourakis**, J. Gao, B. Dolan. A Persona-Based Neural Conversation Model. *ACL*, 2016.

N. Theodorou, **G. Spithourakis** and V. Assimakopoulos. Investment portfolio management with trend forecasting methods. Paper presented at *the Operational Research Society Annual Conference (OR54)*, 2012.

E. Mountalas, **G. Spithourakis** and V. Assimakopoulos. Empirical application, evaluation and comparison of forecasting methods for the electrical system's System Marginal Price: a case study in Greek electricity market. In *Proceedings of the 23rd National Conference of the Hellenic Operational Research Society*, pages 9-13, 2012.

Workshop Proceedings

G. Spithourakis, S. Petersen and S. Riedel. Clinical Text Prediction with Numerically Grounded Conditional Language Models. *EMNLP-LOUHI workshop 2016*.

G. Spithourakis, S. Petersen, and S. Riedel. Harnessing the predictive power of clinical narrative to resolve inconsistencies and omissions in EHRs. *2nd Workshop on Machine Learning for Clinical Data Analysis, Healthcare and Genomics, NIPS 2014*.

Journals

G. Spithourakis, F. Petropoulos, K. Nikolopoulos and V. Assimakopoulos. Amplifying the learning effect via a forecasting and foresight support system. *International Journal of Forecasting*, 31(1):20-32, 2015.

G. Spithourakis, F. Petropoulos, K. Nikolopoulos and V. Assimakopoulos. A systemic view of the ADIDA framework. *IMA Journal of Management Mathematics*, 25(2): 125-137, 2014.

F. Petropoulos, K. Nikolopoulos, **G. Spithourakis** and V. Assimakopoulos. Empirical heuristics for improving intermittent demand forecasting. *Industrial Management & Data Systems*, 113(5):683-696, 2013.

G. Spithourakis, F. Petropoulos, M.Z. Babai, K. Nikolopoulos and V. Assimakopoulos. Improving the performance of popular supply chain forecasting techniques: an empirical investigation. *Supply Chain Forum: an International Journal*, 12(4):16-25, 2012.

Arxiv

B. Riedel, I. Augenstein, **G. Spithourakis** and S. Riedel. A simple but tough-to-beat baseline for the Fake News Challenge stance detection task. *arXiv preprint*, 2017.

Dissertations

G. Spithourakis. Column generation in loopy models. *MSc in Computational Statistics and Machine Learning, UCL*, supervised by Sebastian Riedel, 2013.

G. Spithourakis. Application of aggregate-disaggregate forecasting methodology to non-intermittent demand data. *Diploma (BSc/MSc) in Electrical & Computer Engineering, NTUA*, supervised by Vasilis Assimakopoulos, 2011.

INVITED TALKS

Human Code Computer Tongue: A Human Neural Network. *Welcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS)*, London, UK, March 2018

Human Code Computer Tongue: A Human Neural Network. *UCL Culture*, London, UK, January 2018

Numerically Grounded Language Models. Cambridge NLIP seminar series, *University of Cambridge*, UK, 2017

Text Mining and Sentiment Analysis. Operations Analytics course, *University of Cardiff*, UK, March 2015

SCHOLARSHIPS, GRANTS & AWARDS

Scholarships

- PhD Studentship, The Farr Institute of Health Informatics Research, 2013 – 2016
- MSc Scholarship, Foundation for Education and European Culture (IPEP), 2012 – 2013
- BSc/MSc Scholarship, Estate of Panagiotis Triadafyllidis, 2007 – 2011

Grants & Awards

- Prize (for 3rd place out of 50 in stance detection task), Fake News Challenge, 2017
- Grant (for organising public engagement project), UCL Train and Engage, 2017
- Grant (for travelling to present at conference), EMNLP Student Scholarship Programme, 2016
- Award (for exceptional score in nationwide university entrance exams), President of the Hellenic Republic, 2008
- Award (for 1st place among municipality in nationwide university entrance exams), Eurobank EFG Group ("Great Moment for Education" Programme), 2007

OTHER SKILLS & ACTIVITIES

Languages

Greek (Mother Tongue) • English (Proficient/C2) • German (Advanced/C1) • Japanese (Intermediate/JLPT N3) • Spanish (Intermediate)

Activities

- Organiser of public engagement project "Human Code Computer Tongue" (series of events for teaching about neural networks through poetry generation: performance, December 2017; workshop, October 2017; workshop, February 2017), in collaboration with Zena Edwards and Apples & Snakes, funded by UCL Train and Engage.
- Student volunteer, EMNLP conference, November 2016
- Volunteer in environmental programme, Wales, UK, (2 weeks) 2012
- Volunteer in Japanese-to-Greek translation of Junko Morimoto's short story "Watashi no Hiroshima" ("My Hiroshima"), 2011