Ilya X Valmianski

(619) 977-5721 ivalmian@gmail.com
http://linkedin.com/in/ilya-x-valmianski
http://ilya.valmianski.com

Immigration Status: U.S. Citizen

EDUCATION

Ph.D. Physics 2017
University of California, San Diego
B.S. Biophysics 2009
Summa Cum Laude
University of California, San Diego

INTERESTS

My main interests are in the development of coherently reasoning conversational agents. I think healthcare domain provides both a useful testbed (medical reasoning is hard!) and a high probability of impacting many people. My research focus is on how external knowledge can be leveraged to augment natural language understanding (NLU), generation (NLG), and dialogue management. Here, healthcare is also an excellent domain as labeled data is typically sparse, but there is a rich ecosystem of ontologies and other sources of knowledge.

EXPERTISE

Natural Language Processing
Transformers, recurrent neural networks, n-gram models
Discrete Data Processing
Deep learning with heterogenous data, boosting and other tree models
Machine Learning in Healthcare
Supervised and unsupervised modeling of medical dialogue and clinical progress notes, analysis of discrete EHR data, explainable machine learning model decisions for clinicians

PROGRAMMING

Python ●●●●●
SQL ●●●●●
Matlab ●●●○○
C/C++ ●●○○○

Multilingual fluency: English, Russian
Hobbies: hiking, blues dancing

EXPERIENCE

Curai Health

Staff Machine Learning Researcher 2022-present
Senior Machine Learning Researcher 2021-2022

- Developing natural language understanding and generation models for dialogue contextualization, discourse modeling, and summarization.
- Developing models for clinical decision support recommender systems.
- Supervised multiple graduate-level research interns with the work resulting in 5 papers in the first year (published or in-review).

Kaiser Permanente

Lead Data Architect (Machine Learning) 2019 – 2021

- Lead the development of a symptom checker, SmartTriage, driven by ML models trained on finding diagnoses extracted from ambulatory progress notes and providing clinical decision support to physicians. Delivered the project to pilot stage, after my departure the project was widely deployed in KP Southern California serving millions of encounters per year.

Data Architect (Machine Learning) 2018 – 2019

- Developed a deep learning model for segmenting discourse structure ("sections") in clinical notes. Model deployed to production doing real time inference on >100M clinical notes per year.
- Developed discrete data HCC diagnoses evidence models (predicting thousands of ICD-10 diagnostic codes). Models deployed to production analyzing KP Medicare and ACA patient populations (>1M patients).

University of California, San Diego

Postdoctoral Research Fellow 2017
Graduate Research Assistant (Physics) 2011 – 2017
Graduate Research Assistant (Neuroscience) 2009 – 2011

SELECTED PUBLICATIONS

R. Compton, **I Valmianski**, *et al* "MEDCOD: A Medically-Accurate, Emotive, Diverse, and Controllable Dialog System" *ML4H 2021* https://proceedings.mlr.press/v158/compton21a.html
I Valmianski, *et al* "SmartTriage: A system for personalized patient data capture, documentation generation, and decision support" *ML4H 2021* https://arxiv.org/abs/2010.09905
I Valmianski, *et al* "Evaluating robustness of language models for chief complaint extraction from patient-generated text" *NeurIPS 2019 ML4H Workshop* https://arxiv.org/abs/1911.06915
I Valmianski, *et al* "Microscopy image segmentation tool: robust image data analysis", *Rev. of Sci. Inst.* 85 (3) pp 33701 (2014)
I Valmianski, *et al* "Automatic identification of fluorescently labeled brain cells for rapid functional imaging", *J. Neurophys* 104 (3) pp1803-1811 (2010)

Overall metrics: >25 publications, >560 citations, h-index: 14
Google Scholar: https://scholar.google.com/citations?user=HsOak4YAAAAJ