Dear Deep Learning Summer School Selection Committee,

I'm Dániel Varga, an applied mathematician with a PhD in Natural Language Processing. I've spent my last couple of years in industry as a senior data scientist. Recently I got the opportunity to work at Balázs Szegedy's research group at the Alfréd Rényi Institute of Mathematics of the Hungarian Academy of Sciences. Even our first project is still in an early phase, hence I don't have any directly relevant papers or preprints to share, but the Theano code is public. It implements our novel learning algorithm for generative neural networks. The readme provides some explanation and visualizations:
https://github.com/danielvarga/earth-moving-generative-net
I don't have a traditional personal homepage, so I set this page up to fulfill the requirement.
Click here for a PDF version of this very page.
It's based on my:

- Linkedin page,
- Google Scholar publication list,
- and industry-style CV.

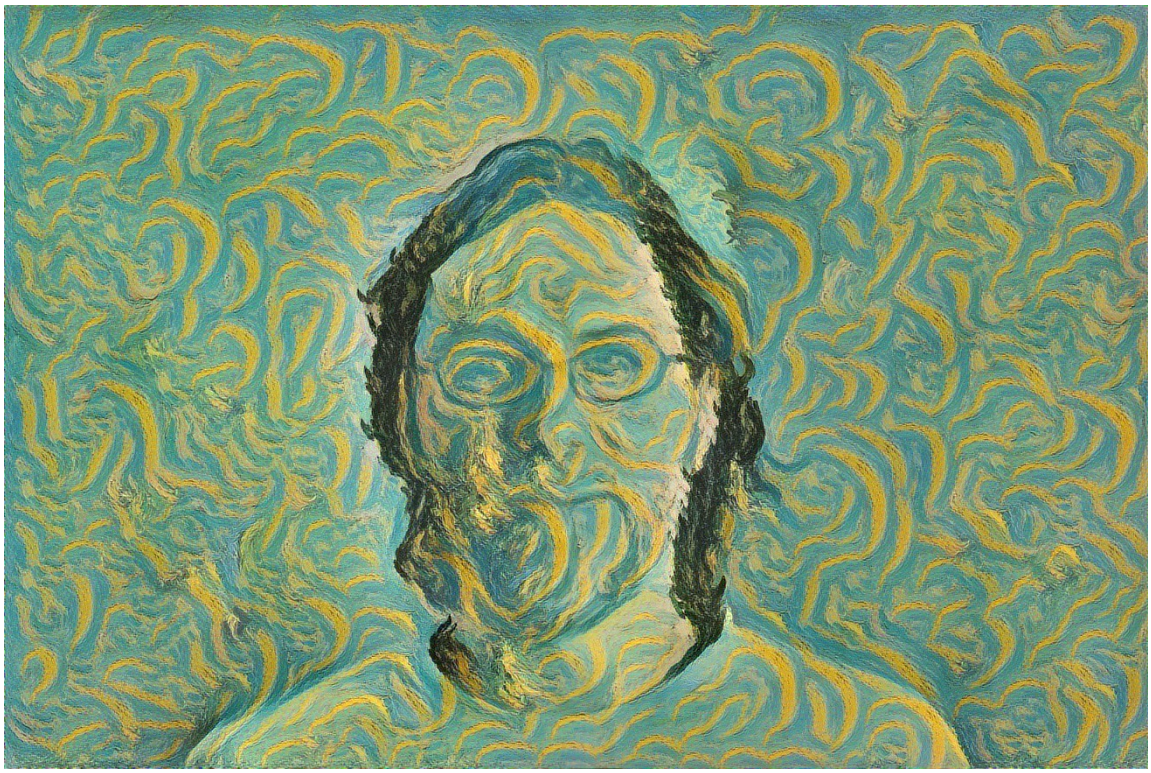

Image created by my own fork of image-analogies, capable of arbitrarily high resolution style transfer.

Click image for my "Cheshire Cat" animation. Visualizes neural network attention based on truncated content weights. Note: 106 MB gif, wait until download finishes.

Research interests

- Generative neural networks. Representational power of small width neural networks.

Selected publications

- Dániel Varga, László Németh, Péter Halácsy, András Kornai, Viktor Trón. Parallel corpora for medium density languages. Recent Advances in Natural Language Processing IV. Selected papers from RANLP 2005. John Benjamins 2007. (222 citations)
- Ralf Steinberger, Bruno Pouliquen, Anna Widiger, Camelia Ignat, Tomaz Erjavec, Dan Tufis, Dániel Varga. The JRC-Acquis: A multilingual aligned parallel corpus with 20+ languages. In the Proceedings of the LREC 2006. (415 citations)
- Viktor Trón, András Kornai, György Gyepesi, László Németh, Péter Halácsy, Dániel Varga. Hunmorph: open source word analysis. Proceedings of the ACL Workshop on Software, 2005. (58 citations)

Education

- PhD in Information Science and Technology, Eötvös Loránd University, Budapest, 2013. Summa cum laude. Thesis: Natural Language Processing of Large Parallel Corpora. Advisor: András Kornai
- MSc in Mathematics, Eötvös Loránd University, Budapest, 1996. Passed with high distinction (red diploma). Thesis: Algebraic Methods in the Theory of Boolean Circuit Complexity. Advisor: Katalin Friedl

Work history

- 2015-present: Research Fellow at the Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences. Member of the Limits of Structures Research Group.
- 2012-present: Senior Data Scientist at Prezi. Information retrieval, machine learning, and modeling-heavy data analysis.
- 2004-2015: Research Fellow (previously Research Assistant) at the Budapest University of

Technology, Department of Sociology, Media Research Centre. Selected projects: hunalign sentence aligner, hunglish.hu searchable English-Hungarian parallel corpus, hunner named entity recognizer, EOVK gerrymandering web application.
- 2003: Researcher at Applied Logic Laboratory Ltd., Budapest. Created SoundStore, a speech indexing system.
- 1997-2002: Researcher at Mindmaker Ltd.

Full list of publications

- Dániel Varga, László Németh, Péter Halácsy, András Kornai, Viktor Trón. Parallel corpora for medium density languages. Recent Advances in Natural Language Processing IV. Selected papers from RANLP 2005. John Benjamins 2007.
- Trón Viktor, Németh László, Halácsy Péter, Kornai András, Gyepesi György, Varga Dániel. Hunmorph: open source word analysis. In the Proceeding of the ACL. ACL 2005.
- Halácsy Péter, Kornai András, Oravecz Csaba, Trón Viktor, Varga Dániel. Using a morphological analyzer in high precision POS tagging of Hungarian. In the Proceedings of the LREC 2006.
- Kornai András, Halácsy Péter, Nagy Viktor, Oravecz Csaba, Trón Viktor, Varga Dániel. Web-based frequency dictionaries for medium density languages. In the Proceedings of the Web as a Corpus workshop at EACL 2006.
- Ralf Steinberger, Bruno Pouliquen, Anna Widiger, Camelia Ignat, Tomaz Erjavec, Dan Tufis, Dániel Varga. The JRC-Acquis: A multilingual aligned parallel corpus with 20+ languages. In the Proceedings of the LREC 2006.
- Farkas Richárd, Simon Eszter, Szarvas György, Varga Dániel. GYDER: Maxent metonymy resolution. Proceedings of the SemEval 2007 Workshop.
- Varga Dániel, Simon Eszter. Hungarian named entity recognition with a maximum entropy approach. Acta Cybernetica Volume 18 Issue 2. 2007.
- Halácsy, P., Kornai, A., Németh, P. & Varga, D. (2008) Parallel creation of gigaword corpora for medium density languages - an interim report, in ‘Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC’08)’, European Language Resources Association (ELRA), Marrakech, Morocco.
- Janky, B., Varga, D. (2013) The poverty-assistance paradox, Economics Letters, Volume 120, Issue 3, pages 447-449, ISSN 0165-1765
- Pléh, Cs., Németh, K., Fazekas, J. & Varga, D. (2011), Entropy measures and predictive recognition as mirrored in gating and lexical decision over multimorphemic Hungarian noun forms, in ‘QMMMD Workshop, University of California, San Diego’.
- Zséder, A., Recski, G., Varga, D. & Kornai, A. (2012), Rapid creation of large-scale corpora and frequency dictionaries, in ‘Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12)’, European Language Resources Association (ELRA), Istanbul, Turkey.
- Csaba Pléh, Kornél Németh, Dániel Varga, Judit Fazekas, and Klára Várhelyi (2013). Entropy measures and predictive recognition as mirrored in gating and lexical decision over multimorphemic Hungarian noun forms. Psihologija, 46(4), 395-418.
- Najeh Hajlaoui, David Kolovratnik, Jaakko Väyrynen, Ralf Steinberger and Daniel Varga (2014), DCEP -Digital Corpus of the European Parliament, in ‘Proceedings of the Nineth International Conference on Language Resources and Evaluation (LREC’14)’, European Language Resources Association (ELRA), Reykjavik, Iceland.