David Pollack

Experience

Co-Founder and CTO @ Cygnus Association Management

(Atlanta — Aug 2014 - Present)

Founded an real estate management company to manage residential and commercial properties in Atlanta, GA. Responsible for the organization's technological solutions, including but not limited to Google Apps infrastructure, company web presence, and payment services. Collaborate with other cofounders regarding all strategic aspects of the company. Work 100% remotely.

Volunteer Program Director @ Rütli Schule

(Berlin — Jan 2014 - Jul 2016)

Lead an optional English program for students in the 11th, 12th, and 13th class at the Rütlischule in Neukölln. We guide students through various web apps that they can use to improve their English, while encouraging students to overcome their fear of speaking with native English speakers.

Senior Consultant @ Navigant

(Atlanta — Sep 2006 - Apr 2013)

Completed re-underwrites of over 750 residential home loans for several of the TBTF (Too Big To Fail) US banks in disputes related to the US housing crisis. Managed and reviewed team members' work product. Developed software to automate repetitive tasks that cut the per-loan prep time by an order of a magnitude. Promoted in 2007 and 2010.

Machine Learning Engineer @ Solvemate GmbH

(Berlin — Feb 2020 - Present)

NLP for chatbots.

Machine Learning Engineer @ i2x GmbH

(Berlin — Apr 2019 - Feb 2020)

I primarily worked on NLP projects at i2x. This included coordinating with a team of over 15 labelers to create datasets from our raw data, creating tooling to manage the data pipeline from raw data to production datasets, and even performing labeling myself (all data was German). For these completed datasets, I wrote tooling to train models and gather metrics with libraries like huggingface transformers (BERT), fasttext, sklearn and others. I integrated pytorch models into our GRPC production, which was previously tensorflow only. Another project that I worked on was a sentence level similarity search using word embeddings and approximate nearest neighbors, tool capable of searching tens of millions of utterances in milliseconds. This project was also dockerized and GRPC-based for easy of deployment. Although my primarily responsibilities were NLP, I experimented with several newer ASR systems (fairseq, wav2letter, and NVIDIA's NeMo) and even contributed several PRs to NeMo. Regarding general software development, code optimizations was a task that I excelled at specifically, I often built optimized dataloaders that could process data online during training, which sped up training and greatly reduced memory requirements for large datasets. I also believe in contributing back to upstream OSS projects and had PRs accepted at huggingface transformers, facebook's FBGEMM, pytorch, pytorch tutorials, and PyThaiNLP. Lastly, due to my proficiency with linux, I was the ML team's de facto systems administrator for our in-house multi-GPU dev machines, so I'm proficient with the *nix cli ecosystem.

Machine Learning Fellow @ Fellowship.ai

(Berlin — Jan 2019 - Apr 2019)

Once again, I developed an end-to-end solution to do image classification on a fashion dataset. We achieved state of the art results in the task classifying a fashion style from the Fashion14 dataset. In preparation for a weekly reading group, I optimized an implementation of memory networks on the babi question answer dataset. Worked on implementing semi-supervised semantic segmentation utilizing Deep Extreme Cut and DeepLab-v3+.

Developer @ PyTorch Audio

(Berlin — Jun 2017 - Jan 2019)

I was the lead developer for PyTorch's official audio loading library, torchaudio, and continue to be involved in the project. I work directly with the PyTorch team at Facebook on this project. I have implemented input error checking testing, audio datasets, variable length input collate functions, audio IO funcationality via SoX (Sound eXchange), and added audio transformations (both implemented directly in PyTorch and using PyTorch's c++ hooks for SoX's effects chain). Additionally, I transferred the main PyTorch project's code flaking and documentation standards to this project. In October 2018, I was invited to the PyTorch developer conference for my contributions to the project.

Machine Learning Engineer @ YEAY GmbH

(Berlin — Jan 2018 - Jun 2018)

My primary role was to use object recognition on videos to identify different classes of clothing. I utilized a Mask R-CNN-based network with custom additions for blur detection to select the clearest frames. Data prep included finding an appropriate dataset, writing scripts to translate between various annotation formats (PASCAL, COCO, YOLO) and add missing fields such as segmentation naively or algorithmically (i.e. GrabCut). Additionally, I dockerized my solution to make it more portable.

Creator @ Joomla! Google Maps Plugin

(Atlanta — Jun 2005 - Sep 2006)

Developed and supported the first Google Maps API plugin for the popular CMS, Joomla! 1.0, in the PHP programming language. Implemented a reverse geocoder prior to the inclusion of this feature in the official Google Maps API. Provided customer support to thousands of users from all over the world. Monetized the project with Google AdWords.

Education

- Master's Program in Economics and Management Science from Humboldt Universität in 2018
- Bachelors of Arts in Economics from Emory University in 2005

Skills

- Proficiency in Deep Learning (pytorch, tensorflow, keras, scipy, etc). Exposure to Scala (Hadoop/Spark) and R.

Github

- Username: dhpollack
- pytorch/pytorch
  - 26168 - allow building docker without torchvision
  - 3505 - improvements to ModuleList and ParameterList classes
  - 2657 - Add PadNd functions
- pytorch/audio
  - 72 - MEL2 optimzation, filter bank conversion fix, flake8 refactor
  - 59 - allow loading with offsets and num_frames and saving specified bit precision
  - 58 - torch.cos not implemented for LongTensor fix
  - 23 - pytorch implementation of MEL spectrograms (no librosa req'd)
  - 16 - mu-law companding transform
  - 15 - add MEL spectrograms transform and fixed a few tests
  - 13 - update save/load to prevent bad inputs
  - 11 - basic transforms and two datasets
  - 10 - add gitignore
  - 4 - bytes function requires type
- huggingface_transformers
  - 907 - Fix convert to tf
  - 1089 - change layernorm code to pytorch's native layer norm
- nvidia_nemo
  - 38 - refactor Manifest class and some optimizations

    - 28 - allow for labels with length > 1
- fluxjl
    - 634 - instance normalization
- rasa_nlu
    - 443 - add sklearn-crfsuite install instructions
- more...
    - 0 - my PRs

Languages

- English (native)
- German (professional working proficiency)
- French (basic)

Activities

- Duolingo
- Coursera
- MoBro @ Movember
- Cycling
- Bouldering

| Braunschweiger Straße 54, 12055 Berlin | +49 152 0476 3223 | david.pollack@cms.hu-berlin.de |