Pratik Deoolwadikar

pratik.deoolwadikar@gmail.com +91 9769118579

WORK EXPERIENCE

Data Scientist, Netcore

Apr 2019 - Aug 2020

- Developed Email Auto-tagging system utilizing Knowledge Graphs, realtime information retrieval techniques and vector Clustering algorithms to obtain semantic categorization.
- Explored performance of Attention based pretrained Encoder-Decoder Transformer language models like BERT, Universal Sentence Encoder etc and few Bidirectional approaches on various NLU, NLG applications like text generation and chatbot.
- Wrote novel Clustering/Parser algorithm to obtain custom clusters from time series data frames and crafted features using big analytical queries on analytical db.

Data Scientist, NanoPrecise Sci Corp

Nov 2019 - Apr 2020

- Developed health monitoring and fault detection system for industrial assets utilizing Signal Processing, Math modeling and Sequential Machine Learning algorithms.
- Analyzed acoustic data from sensors employing domain based noise filtering to discover failure patterns. Designed data labeling pipelines for online learning.
- Designed bootstrap algorithm for battery life estimation of sensors. Explored algorithm performance with changes to spatial positioning and intermittent historical utilization.
- Designed, built the improved data pipelines and backend architecture utilizing Multi-Processing and Virtualization along with AWS S3, Redis, Apache Kafka, Docker etc. frameworks.

PROJECTS

Developed PyPI Framework, Transformers Keras Dataloader

- Enables real-time data feed to Transformer models for downstream training, unlocking the capacity to handle bigger datasets and larger batch sizes.
- Provides support to utilize GPU and Multi-Processing for input processing and computation.
- Added support for custom layer pooling strategies to generate word/sentence input vectors.

Scratch python implementations of Machine Learning algorithms

- Multi flavoured, multi-variate and multi-nomial implementations for 12 foundational types of machine learning algorithms, in an effort to understand mathematics and architecture backing these algorithms.
- Wrote scratch implementations for many known Optimizers, Activations, Initializers.

Transliteration using Encoder-Decoder Attention model

- Pytorch implemention of Encoder-Decoder Attention model to transliterate text from source(hindi) to target(english) script.

PyTorch implementations of Deep Learning algorithms

- Explored and implemented various flavours for popular Deep learning algorithms like CNN, RNN, LSTM, Encoder-Decoder, etc in PyTorch.

Slot filling using CRF and BiLSTM

- Entity slot identification in BIO format using Conditional Random Fields as slot filter on top of BiLSTM to adapt slot dynamics of training corpus.

Neural Relation Extraction using pretrained Language model

- Semantic relation extraction of marked entities from documents, utilizing language model for obtaining word/phrase representations and downstream classifier to map entity pair similarity to all possible relations.

Finetuning Transformer Language models

- Finetune pretrained parameters of Transformer Language models for text classification task, written a super-fast solution by employing techniques like gradient accumulation, dynamic padding, smart batching and mixed precision.

Multi-Armed Bandit Problem

- Studied various reinforcement learning approaches of exploration and exploitation to solve K-armed bandit problem.

Human Activity Recognition, LSTM on TensorFlow Android

- Realtime activity prediction from continuous spatial data of Accelerometer on Android, to classify amongst six different human activities.

Audience Segmentation, Graph Neural Network

- Segment audiences by categorizing complex relationships using GNN, trained on engineered features from email corpus.

Semantic Topic Clustering, Universal Sentence Encoder

- Clustering topics using attention based language model, to group based on semantic relationships among subject topics for analysis.

Behavioural Cloning, Convolutional Neural Network

- Used CNN to predict steering angle from augmented first person images of road & scene to drive a car in the simulator, as a part of Udacity Self Driving Car nanodegree.

SKILLS

Programming Languages

Python, Java, C++, JavaScript, C, HTML/CSS, C#, PHP & MySQL etc.

Frameworks

PyTorch, TensorFlow, Keras, Android framework, NodeJS, Redis, Docker, Vertica DB, MongoDB, Apache Spark, AWS3, EC2, Django, Rails, Matlab, MLFlow, HuggingFace, Tensorboard etc.

EDUCATION

B.E in Computer Engineering

Aug 2018

A.P. Shah Institute of Technology, D.T.E University of Mumbai, India

Diploma in Mechanical Engineering

Jun 2015

Maharashtra State Board of Technical Education, Thane, India

CERTIFICATION

Machine Learning Engineer, Udacity Aug 2018

- Hands on projects using Machine Learning, Deep Learning, Reinforcement Learning.
- Elementary projects from NLP, Computer Vision/Image Processing etc.
- Exposure to ML/DL frameworks like Tensorflow, Keras, Pytorch etc.

AWARDS

Smart India Hackathon 2017

Awarded by:

- Ministry of Road Transport and Highways, Government of India.
- Persistent Systems Ltd.

LINKS

Github

github.com/pratikdk

Website

pratikdk.github.io

LinkedIn

linkedin.com/in/pratikdeoolwadikar