Arnab Pushilal

Skilled data scientist with 3 years of experience of developing data solutions for the automotive industry. I like to learn and write about what I am currently thinking about. I'm interested in deep learning, NLP, vision, and developing machine learning-based applications. Presently, I am working on Recommender Systems that are health-aware. Please do visit my website at the link below for a detailed bakground.

✉ thestuffedshow1@gmail.com
📍 London, United Kingdom
⌨ github.com/ArnabPushilal
📱 +44-7880205911
🖱 arnabpushilal.github.io/
◑ medium.com/@arnabpushilal

EDUCATION

BTech Electrical & Electronics Engineering
VIT University, Vellore

05/2014 - 05/2018   *CGPA: 9.07/10 - Top 5%ile*

Relevant Modules

- DSA, Neural Networks, Numerical Methods, C programming, Signal Processing
- Thesis: Facial Recognition for Home Security Systems, Published Article ID: 20706, IJET

MSc Data Science & Machine Learning
University College London

09/2021 - Present   *Average 86.2%*

Term-1/Term-2 Modules

- Intro to ML (90), Statistical Data Science (90.42), Deep learning (82.5), Machine Vision (75.10)
- Applied ML (95),Data Mining (88.5), AI for sustainable development (86.5), NLP (81.8)

WORK EXPERIENCE

Asst Manager, Data Acquisition & Analysis
Hero MotoCorp

07/2018 - 06/2021

Achievements/Tasks

- Designed a model to predict Emission failure from sensor data captured from two-wheelers in a lab setting that effectively saved 1 million pounds for the company & was deployed to 10 locations across the country.
- Wrote scripts to automate data visualization and reporting. Users could login their data & the charts would automatically update and generate reports periodically.
- Led the data-devolopment team who were responsible for collecting & reporting the measurements & experiments.

Tech Stack : Python, PyTorch, git, Flask, numpy, Pandas

Internship
ZOE

06/2022 - Present   *London, United Kingdom*

Project

- Working on recommeding healthier recipes to ZOE user's combining user-preferences with nutritional value.
- Devoloped a new ingredient-ingredient self-attention + user-ingredient attention model that increased the AUC on the task by 5%.
- Used BERT sentence embeddings to cluster recipes and find healthier versions of the same recipes.

Tech Stack : Python, Pytorch, git, numpy, pandas, GCP

SKILLS

Python · SQL · Pytorch · Pandas · Numpy · Sci-kit · Flask · Tensorflow · git · PySpark · NLTK/Spacy · Docker · Linux/Windows · GCP

PROJECTS

whatsmyageagain.herokuapp.com (06/2022) ⧉
- Devoloped a age-prediction web app using a face detector and a VGG-based age prediction model.
- Used Flask, SQLalchemy, HTML to create front-end & back-end, used docker & Heroku to deploy.

Automatic Neutralization of Sexist language (02/2022) ⧉
- Found novel ways of evaluating neutralized sentences, contributed a new dataset for the task.

Multi-task Learning using attention & auxillary tasks (11/2021) ⧉
- Trained a attention-based model to with self-supervised auxillary tasks like canny-filter edge detection, denoising, colourization to improve target task of segmentation.

Binary Classifier with Mixture of Gaussians (10/2021) ⧉
- Trained a MOG model using E-M for background and target pixels and calculated the posterior to predict the probability, and F1, ROC to evaluate the model.

Spotify Popularity Predictor (09/2020) ⧉
- Collected data from Spotify API & used various models to predict the Popularity Score for any song.

CERTIFICATES

Coursera: Tensorflow Specialization in Practice, IBM Data Science Professional

Udemy: PySpark for Bigdata.

ARTICLES FOR ANALYTICS VIDHYA

Understanding transformers ⧉

Character aware CNNs ⧉

Word Embeddings ⧉

Tokenizers ⧉

Recommender Systems: Overview ⧉

Kernels ⧉