

# Arnab Pushilal

Skilled data scientist with 3 years of experience of developing machine learning solutions for the automotive industry. 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.

✉ thestuffedshow1@gmail.com

📍 London, United Kingdom

🐙 github.com/ArnabPushilal

📞 +44-7880205911

📄 arnabpushilal.github.io/

🗣 medium.com/@arnabpushilal

## WORK EXPERIENCE

### Machine Learning Engineer Intern

ZOE

06/2022 - Present

London, United Kingdom

Project

- Built a **recommendation engine** to suggest healthier recipes to ZOE user's combining user-preferences with nutritional value.
- Developed a novel **ingredient-ingredient self-attention + user-ingredient attention model** that increased the AUC on the task by 5%.
- Evaluated **BERT sentence embeddings** to cluster recipes and find healthier versions of the same recipes.

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

### Data Analyst

Hero MotoCorp

07/2018 - 06/2021

Achievements/Tasks

- Designed a model using **RNNs** to **predict Emission failure** from sensor data captured from two-wheelers in a lab setting. I oversaw this project from the data collection, validation to deployment.
- Wrote scripts to **automate production data visualization** and reporting Emission Measurement. Riders could login, & their ride data & the charts would automatically update and **generate reports periodically**. This reduced the overall lead-time to on fault detection & correction for two-wheelers in production.
- Led the data-development team who were responsible for collecting & reporting the measurements & experiments.

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

## EDUCATION

### BTech Electrical & Electronics Engineering

VIT University, Vellore

2014 - 2018

CGPA: 9.07/10 - Top 5%ile

### MSc Data Science & Machine Learning

University College London

2021 - Present

Module Average 86.2%

Courses

- Machine Learning, Deep Learning, Machine Vision, Statistical Data Science
- NLP, Information Retrieval and Data Mining, AI for sustainable development, Applied ML

## SKILLS

Python

SQL

Pytorch

Pandas

Numpy

Sci-kit

Flask

Tensorflow

git

PySpark

NLTK/Spacy

Docker

Linux/Windows

GCP

## PROJECTS

whatsmyageagain.herokuapp.com (06/2022) [🔗](#)

- Designed a **age-prediction web app** using a face detector and a VGG-based age prediction model.
- **Flask, SQLAlchemy, HTML** to create front-end & back-end, used **docker & Heroku** to deploy.

Automatic Neutralization of Sexist language (02/2022) [🔗](#)

- Utilized a **BERT encoder & an LSTM decoder**. Found a novel method of evaluating neutralized sentences, contributed a new dataset for the task.

Multi-task Learning using attention & auxillary tasks (11/2021) [🔗](#)

- Trained an **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.

## ARTICLES & CERTIFICATIONS

Understanding transformers [🔗](#)

Character aware CNNs [🔗](#)

Word Embeddings [🔗](#)

Tokenizers [🔗](#)

Recommender Systems: Overview [🔗](#)

Kernels [🔗](#)

Coursera: IBM Data Science Professional, Tensorflow specialization

Udemy: PySpark for BigData