# SHASHWAT UPADHYAY

07880313423 • work.shashwat@gmail.com • linkedin.com/in/ushashwat • github.com/ushashwat

## WORK EXPERIENCE

**BT Group** Bristol, United Kingdom

***Digital Graduate – Data Science, Analytics, and Enterprise*** Sep 2022 – present

- Evaluated customer datasets in BigQuery with up to 340 million data points to analyse EE broadband and mobile usage and facilitate strategic cross-selling and upselling of products
- Contributed to a data product that is now being used as a benchmark and has fed upstream marketing systems, delivering around £1.6 million from broadband sales
- Explored broadband abandoned basket datasets on EE web store to track product navigation journey of users and implemented ways towards user retention for more than 70 million web page hits
- Enhanced data quality by over 35% using SQL queries to combine tables across different business units within BT, thus improving modelling pipeline efficiency by almost 20%
- Played a key role in charting monthly cost estimations using 160+ KPIs and automating 90% of the virtual mailbox approval process within Corporate Unit through Power Automate and Power Apps

**Etcembly Ltd.** Bristol, United Kingdom

***Project Member*** Jan 2022 – May 2022

- Decoded the immune response using T-cell receptors for 650K cells & 25K expression levels across 130 patients using the Scanpy toolkit
- Transformed the single-cell sequencing dataset and applied gene filtering methods using highly variable genes and differentially expressed genes for 100-160 top genes and 5 T-cells
- Extracted differential features between healthy & affected patients via a machine learning pipeline to predict the severity and determine the most important transcriptomic features

**Wikilimo** Prayagraj, India

***ML Engineering Intern*** Aug 2020 – Nov 2020

- Used AI-driven climate informatics to predict Normalised Burn Ratio and Green Normalised Difference Vegetation Index on a custom Sentinel Hub Earth Observatory webpage
- Generated custom scripts in Python to extract and load geospatial data from 12 spectral bands within EOPatch into the modelling pipeline
- Processed satellite weather data via ARIMA model with 0.018 as MPE and LGBM regressor with 91.8% as recall score

**Indian Institute of Technology** Indore, India

***AI Intern*** Jun 2019 – Jul 2019

- Worked on 3D ear data provided by the University of Notre Dame for Ear Recognition in Biometrics via PointNet and Siamese network
- Utilised a slicing approach for each axis to inspect and feed 3-dimensional point cloud data to the deep learning model built in Keras and improve the overall performance metrics by 36%
- Attained an accuracy of 86.06% after successfully applying an augmentation technique to the point cloud methodology

## EDUCATION

**University of Bristol** Bristol, United Kingdom

***Master of Science – Data Science***, Award: Distinction Sep 2021 - Sep 2022

Relevant Coursework – Statistical Computing & Empirical Methods, Data Engineering, A.I., Advanced Data Analytics

Dissertation – Empirical analysis of forecasting techniques for prices & returns of tradable financial assets

- Compared the performance of 2 time-series models for capturing trends, seasonality, etc., with 8 ML algorithms
- Tweaked the SGD regressor model to achieve 0.988 as the R-Squared score for next-day rolling forecast

**Dr A.P.J. Abdul Kalam Technical University** Ghaziabad, India

***Bachelor of Technology – Information Technology***, Award: First-Class Honours July 2016 - Sep 2020

Bachelor's Thesis – Medical Image Analysis

- Implemented 6 distinct deep learning models to analyse and quantify the available bio-medical data
- Surveyed Bone Abnormality, Brain Tumour, and Lung Cancer detection with 83% as the highest target score

## PUBLICATION

Fake news detection using deep learning models: A novel approach

- Tested 7 deep learning architectures, including an ensemble of Bi-directional LSTM with an attention mechanism
- Achieved state-of-the-art accuracy of 88.78% in classifying misinformation in tweets on several topics and effectively improving on the previous approach having a detection rate of 85%
- Accumulated over 168 citations for this peer-reviewed research paper, thus helping reduce bias on social media

## SKILLS

**Machine Learning / Data Science –** Pandas, Numpy, Matplotlib, Seaborn, Sci-kit learn, Statsmodels, Tidyverse, Keras, PyTorch, Neural Networks, Image Processing, Natural Language Processing, Data Analysis, Statistical Analysis, Data Visualisation, Cloud Computing, Machine Learning, Deep Learning

**Programming Languages –** Python, SQL, R

**Tools and Technologies –** Git, Tableau, Qlik Sense, Google BigQuery, AWS, Power Automate, Power Apps, JIRA, Microsoft Excel, LaTeX

## RECOGNITION

**BT Generative AI Session –** Presented Gen AI to over 90 colleagues under Digital & Business for corporate use cases

**PyTorch Scholarship –** Awarded a scholarship from Udacity to study PyTorch concepts in production development

## VOLUNTEERING

**Be My Eyes –** Volunteering for 620K+ blind and low-vision people to provide visual assistance through a live video call

**Home-Start (Bristol, UK) –** Participated in raising charitable donations to help families going through challenging times

**Volunteering & Fundraising Network (Bristol, UK) –** Promoted fundraising initiatives for 7600+ postgraduate students

**Slum Swaraj Foundation (Ghaziabad, India) –** Provided free education to around 40 underprivileged children