

# Bence Csiba

[linkedin.com/in/bencecsiba](https://www.linkedin.com/in/bencecsiba) | 07496063113 | [bencecsiba5@gmail.com](mailto:bencecsiba5@gmail.com) | [bencecsiba.github.io](https://bencecsiba.github.io)

## EDUCATION

---

### University of Bath

*BSc (Hons) Physics with Year Long Work Placement*

*Oct 2018 - Jul 2022*

- Awarded Upper Second Class Honours, with an overall average of 66.75%
- Highlights include 80% in Mathematical Methods for Physics 2 (Linear Algebra, Statistical Modelling), 77% in Quantum & Atomic Physics (Probability).
- Achieved a First for my Final Year Project, where I worked in a small team with the Met Office to build a neural network to predict water wave heights from a nearby weather radar using TensorFlow and Py-ART.

### Other Achievements

- Awarded an AWS Scholarship to study the AI programming with Python Nanodegree via Udacity (2023)
  - Built deep learning models, such as a flower classifier using transfer learning from VGG16 in PyTorch
  - Selected for the Advanced Program (Top 25% of students), building models on AWS using PyTorch
- Deep Learning Specialisation, deeplearning.ai via Coursera (2020)
- UKMT Senior Team and Individual (Silver) Maths Challenge (2018)

## RELEVANT EXPERIENCE

---

### Kubrick Group

*Machine Learning Engineer*

*Apr 2023 - Present*

- Completed a comprehensive 17-week training period, covering SQL, Docker, MLOps, Time Series Analysis, Spark, PyTorch, NLP, Deep Learning and command line environments. Also achieved the AZ-900 certification.
- Developed a Random Forest model for predicting banking churn on an artificially generated dataset. Data was analysed to perform feature engineering, highlight deficiencies and to help identify possible remedies. Solution was evaluated using an RoC curve and findings were presented to stakeholders with visual aids generated using PowerBI
- Deployed solutions using FastAPI and Postman and containerised using Docker.
- Collaborated in small teams using git, GitLab and MS Azure.
- Learnt new technologies, including earning the Dataiku Associate Developer certificate and building LLM agents using LangChain

### NextWave Consulting

*Associate Consultant (Appian Developer)*

*Sep 2022 - Apr 2023*

- Achieved the Appian Associate Developer Certification.
- Lead the development of an internal onboarding tool. This included building a coherent and efficient data model, which was built using MySQL. Data was used dynamically throughout to optimise efficiency and reduce the likelihood of mistakes.
- Placed at a FTSE 100 insurer, building an internal tool to automate processes and data pipelines. This allowed data to move between previously disconnected tools, reducing the need to enter data multiple times, thus reducing the time wasted by underwriting assistants.
- Progress was communicated to non-technical stakeholders, including senior figures, business analysts and end users. Data was ingested from user inputs and databases using Kafka.
- Worked as part of an Agile team, using Jira to track progress. Development occurred across different environments for effective version control and the secure deployment of packages and updates

### FedEx Express International

*Project Management Office (Global IT Infrastructure) Intern*

*Oct 2020 - Sep 2021*

- Developed my knowledge of IT infrastructure and how it is managed and upgraded in an efficient manner.
- Analysed data, performed data discovery, and presented findings using Excel and other tools for 100s of servers and 1000s of devices/users worldwide.
- Communicated with IT teams in 100+ countries and with individual users as problems arose.

## OTHER EXPERIENCE

---

### Bath Ultimate

#### Treasurer

2021/2022

- Responsible for the financial running of the club and player participation in large tournaments
- Managed a club budget of £4,000 in addition to money raising activities to support the club
- Analysed and forecasted spending requirements to allow subsidies to encourage people from lower socioeconomic backgrounds to participate

#### Coach

2019/2020

- Responsible for running training sessions for upwards of 50 people, utilising my leadership, organisation, and communication skills. Culminated in a 4<sup>th</sup> place finish at University Mixed Indoor Nationals, 2020

## PERSONAL PROJECTS

---

### FPL Points Predictor

- Created an ETL Process to consolidate data from multiple sources.
- Performed exploratory data analysis and consequently engineered features to optimise model performances. This included temporal windows of varying lengths to capture longer term data.
- Trained a random forest regressor due to the data being in a tabular format. Found optimum hyperparameters using randomised search. Utilised the resulting feature importances to inform decisions regarding splitting and training two separate models for better performance.
- Built a fully connected neural network using PyTorch, using Batch Normalisation and Dropout layers to avoid overfitting. Optimised hyperparameters to obtain best performance
- Utilised the Autogluon tabular predictor to fit a variety of models, the best of which outperformed previous models.
- Used a variety of data structures to find predicted best team and the FPL API to compare the performance of my model with that of global users.

### Skin Lesion Classifier

- Built a neural network to classify pigmented skin lesions in the Skin Cancer MNIST: HAM10000 dataset
- Variance in number of classes, which I established whilst performing exploratory data analysis, was treated using the imbalanced dataset sampler
- Utilised transfer learning from Resnet18 to gain a higher accuracy.
- Evaluated performance of the model using an accuracy score, as well as a confusion matrix to establish performance across classes

### Landmark Classifier

- Final project as part of Udacity Nanodegree, tasked with classifying 50 landmarks
- Built a neural network from scratch, utilising residual blocks to achieve an accuracy score of 51%
- Using transfer learning from VGG16, achieved an accuracy of 74%
- Developed the model on Sagemaker, using S3 buckets to store the data. The model was then deployed to an AWS endpoint for inference.

## SKILLS AND INTERESTS

---

### Skills

- Python: Tensorflow, PyTorch, Pandas, Scikit-learn, Keras
- SQL, Docker, Appian, Hungarian (Fluent)

### Ultimate

- Selected to represent Hungary at the 2023 European Championships
- Competed in club-level competitions in the UK and in Europe

### Music

- Keen pianist and flautist, having achieved Grade 8 in both (Merit and Distinction respectively)