FINAL-YEAR PROJECTS

Supervisor: Barry Dillon

1. **Project Title:** Fast/efficient MatMul-free machine-learning
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning, transformer networks, physics, linear algebra
   Proposed Tech Stack: Python, PyTorch, SciKit-Learn, Jupyter, Google Colab
   Summary: The application of machine-learning models has seen a massive growth in recent years, however the models being employed are also growing in size. There is a need for machine-learning models which can perform much of the same tasks, but with a much smaller footprint. In this project the student will study the use of MatMul-free neural networks for classification tasks at the Large Hadron Collider experiment. These networks have a much smaller memory and power footprint than standard neural networks. No previous experience or knowledge of physics or particle physics is required, but some will be obtained during the research. Clean datasets will be provided for the project.
   Reference: https://arxiv.org/abs/2406.02528v1

2. **Project Title:** Anomaly detection in experimental physics data
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning, neural networks, anomaly detection, physics
   Proposed Tech Stack: Python, PyTorch, SciKit-Learn, Jupyter, Google Colab
   Summary: Anomaly detection is a machine-learning technique designed to find anomalous points in a dataset. There are many ways to define what is anomalous in a dataset. Some of the techniques are completely unsupervised, for example AutoEncoders, and others are semi-supervised and use some labelled data. In this project the student will study anomaly detection methods on simulated event data from a physics experiment at the Large Hadron Collider. No previous experience or knowledge of physics is required for the project. Clean datasets will be provided for the project.
   Reference: https://atlas.cern/Updates/Briefing/Anomaly-Detection

3. **Project Title:** Object anomaly detection in experimental physics data
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning, neural networks, anomaly detection, physics
   Proposed Tech Stack: Python, PyTorch, SciKit-Learn, Jupyter, Google Colab
   Summary: Anomaly detection is a machine-learning technique designed to find anomalous points in a dataset. There are many ways to define what is anomalous in a dataset. Some of the techniques are completely unsupervised, for example neural networks, and others are semi-supervised and use some labelled data. In this project the student will study anomaly detection methods on simulated jet data from a physics experiment at the Large Hadron Collider. The jet data focuses on specific objects, jets, measured in proton collisions at the experiment. No previous experience or knowledge of physics is required for the project. Clean datasets will be provided for the project.
   Reference: https://atlas.cern/Updates/Briefing/Anomaly-Detection

4. **Project Title:** Solving differential equations with neural networks

Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
Key Topics: Machine-learning, basic differential equations, neural networks
Proposed Tech Stack: Python, PyTorch, SciKit-Learn, Jupyter, Google Colab
Summary: Differential equations are be used to describe how variables evolve as some other parameters change, for example time. While simple differential equations can be solved exactly, more complex non-linear equations need approximate numerical solutions. Traditionally, techniques such as the Runge-Kutta have been used to find these solutions. More recently however, neural networks have been showing to be very effective at finding these solutions also. In this project the student will study the application of neural networks to solve systems of differential equations that arise in a framework used in high-energy physics called Effective Field Theories. No previous experience or knowledge of physics is required. Some knowledge of differential equations is required.

5. **Project Title:** Understanding scientific data with topic models
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning, Bayesian modelling, topic models, physics
   Proposed Tech Stack: Python, SciKit Learn, Jupyter, Google Colab
   Summary: Data collected in physics experiments can arise from a multitude of different sources. Disentangling these sources in the data is important for analyses. In this project the student will use Bayesian topic models such as Latent Dirichlet Allocation to analyse simulated data from the Large Hadron Collider experiment, disentangling the effects in the data from number of different processes. The Bayesian methods work by identifying co-occurring features in the data and extracting patterns in an unsupervised manner. No previous experience or knowledge of physics or particle physics is required for the project. Clean datasets will be provided for the project.

6. **Project Title:** Classification of objects and events in physics experiments
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning, neural networks, physics
   Proposed Tech Stack: Python, PyTorch, SciKit-Learn, Jupyter, Google Colab
   Summary: Data collected in experiments can be difficult to classify, and can be too large to classify by inspection. This is particularly the case for the Large Hadron Collider experiment. To overcome this, the use of automated machine-learning tools has become very popular. In this project the student will explore the use of supervised machine-learning techniques to classify signals from rare processes in simulated data. No previous experience or knowledge of physics is required for the project. Clean datasets will be provided for the project.

7. **Project Title:** Regression techniques for observatory data
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning, neural networks, physics
   Proposed Tech Stack: Python, PyTorch, SciKit-Learn, Jupyter, Google Colab
   Summary: Regression techniques allow us to model relationships between different features in a system. This can be useful in scientific experiments, where some features are easier to measure than others, and you can then use the regression model to predict those that are unable to be measured from those that can be measured. This is the case in observatories measuring the particles scattered by cosmic rays in the upper

atmosphere. In this project the student will use machine-learning based regression techniques to predict the depth of cosmic-ray induced air showers from ground-based observations. No previous experience or knowledge of physics is required for this project. Clean datasets will be provided for the project.

8. **Project Title:** Open Project – anomaly detection
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning
   Proposed Tech Stack: Python, PyTorch, SciKit Learn, Jupyter, Google Colab
   Summary: I am willing to supervise projects in the research areas of machine-learning using anomaly detection tools. The student should have one or more datasets in mind when proposing a project.

9. **Project Title:** Open Project – efficient machine learning
   Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
   Key Topics: Machine-learning
   Proposed Tech Stack: Python, PyTorch, SciKit Learn, Jupyter, Google Colab
   Summary: I am willing to supervise projects in the research areas of machine-learning using more efficient machine-learning models, such as MatMul-free models. The student should have one or more datasets in mind when proposing a project.
   Reference: https://arxiv.org/abs/2406.02528v1

10. **Project Title:** Open Project
    Eligible Course: BSc Hons Computer Science, BEng Hons Artificial Intelligence
    Key Topics: Machine-learning
    Proposed Tech Stack: Python, PyTorch, SciKit Learn, Jupyter, Google Colab
    Summary: I am willing to supervise projects in the research areas of machine-learning. When proposing a project, the student should have a very specific idea in mind for what they want to work on, along with one or more datasets.