**Henrique Sousa**
https://hbvsa.github.io/
Email : henriquebvsousa@gmail.com
Phone: +351 960309500

## Skills

- **Technical Skills**: Deep Learning, Machine Learning, Data Science
- **Programming Languages**: Python, SQL, Java, C++
- **Programming Frameworks**: PyTorch, Tensorflow, Keras, HuggingFace, Pandas, OpenCV, Matplotlib
- **Technologies**: Git, Docker and model deployment to cloud platforms

## Experience/Projects

- **Master Thesis: Deep Learning for SNP-Disease Correlation Analysis** Instituto Superior Técnico
  *Deep Learning Research* *September 2020 - October 2022*
  - **Problem and Motivation**: Developed a deep learning model capable of learning and identifying correlations between Single Nucleotide Polymorphisms (genetic data) and probability of disease. The model task is hard since it required learning correlations in datasets with hundreds of features where only two or three features were significant, and sometimes only through purely nonlinear interactions. A short but more thorough description is available at Master Thesis.
  - **Deep Learning Expertise**: Gained in depth intuitive understanding of state-of-the-art deep learning architectures (Transformers, CNNs, RNNs) through extensive literature review and applied the most recent techniques for optimizing the model training.
  - **Custom Model Development**: The model architecture was conceptualized and coded from scratch in Python with layers designed for model interpretability since the primary goal was to identify the significant genetic features after model training.
  - **Significant Results**: The proposed model achieved a 99% detection rate on datasets containing 3 significant features with purely nonlinear interactions (hardest datasets) where the top performing state of the art methods achieved 0% detection rate given their inadequacy for pure nonlinear interactions. The model was also capable of identifying subtle correlations where genetic data influenced only 1% of the disease probability variation.
- **Deep Learning Projects**
  *Deep Learning, Software Engineering* *Nov 2022 - Present*
  - **Deep Learning Projects**: Implemented multiple deep learning models to different projects, especially by reusing SOTA architectures such as ViT(2021). One such project using a SOTA object detection and segmentation vision transformer is showcased at Portfolio. I've gained practical experience in deploying such models using Flask/Gradio as the front-end interface to cloud platforms (Azure, Google Cloud, Amazon) and HuggingFace using Docker.
  - **LLMs**: Have learned to use popular packages such as HuggingFace and techniques to build LLM rag systems, fine-tune LLMs to specific tasks as well as experience with prompting. Some examples are showcased at Portfolio.
- **Quantitative Research and Algorithmic Trading**
  *Software Engineering, Data Science* *Nov 2022 - Present*
  - **Strategy Backtester**: Designed and implemented a comprehensive algorithmic trading backtesting system in Python to perform a complex and rigorous simulation, incorporating data analysis of the historical data of 2000+ financial instruments for more than 30 years.
  - **Data Analysis and Optimization**: Developed a trading strategy which was evaluated using data analysis and proved its reliability and significance using statistical testing.
  - **Visualization Tools**: Developed and integrated visualization tools using Finplot and PyQt libraries for real-time and historical data analysis. These tools proved essential for strategy evaluation, adjustment, and execution debugging.

## Education

- **Instituto Superior Técnico** Lisboa, Portugal
  *Master of Science in Computer Science and Engineering* *2018 – 2022*
  - **Relevant Courses**: Data Science, Machine Learning, Natural Language Processing, Reinforcement Learning, Artificial Intelligence in Games, Autonomous Agents and Multi-Agent Systems
- **ISCTE-IUL** Lisboa, Portugal
  *Bachelor of Engineering in Computer Science and Engineering* *2015 – 2018*