Andrew Mayes

www.linkedin.com/in/andrew-mayes-2763a716a

andrewmayes14@hotmail.com

+61489912397

Machine Learning Engineer at re:mind

Sydney, New South Wales, Australia

About

Machine Learning Engineer specialising in training, fine-tuning, and deploying deep learning models, including **Large Language Models (LLMs)**, **Vision-Language Models (VLMs)**, **Vision Transformers (ViT)**, and **Embedding Models**, to derive actionable insights and produce impactful business solutions. Skilled in building and scaling end-to-end machine learning pipelines using **Docker** for containerisation, **AWS Lambda**, **S3**, and **PostgreSQL** for seamless integration, and robust APIs with **FastAPI** to enable secure and efficient access to model outputs. Passionate about delivering real-world AI solutions, including automating data labeling and validation, transforming unstructured data into structured formats, and streamlining document extraction to optimise business processes. A proactive self-starter and collaborative team player who thrives in cross-functional, multi-domain projects, continuously seeking innovative approaches to drive value across various fields.

Publication

Mayes, A. and Anwar, A., 2022. Machine Learning Based IDS for Cyberattack Classification. In Explainable Artificial Intelligence for Cyber Security: Next Generation Artificial Intelligence (pp. 93-111). Cham: Springer International Publishing.

Experience

Machine Learning Engineer

re:mind | March 2024 – Present | France (Remote)

Focused on transforming unstructured document data into actionable insights by leveraging **LLMs**, **ViT**, **VLM**, and **OCR technology**. Created end-to-end automated solutions to streamline document processing, enhance user experience, and drive critical business outcomes.

Key Achievements:

- **Automated Data Collection:** Built an efficient web scraping pipeline using Google Lens, automating data acquisition and minimising manual input, resulting in faster, more accurate data collection.
- **Conversational Diagnosis:** Designed a contextualised chatbot leveraging **LLMs** with **Retrieval-Augmented Generation (RAG)** for personalised responses based on specific user and product contexts, significantly improving user interaction and satisfaction.
- **Document Type Classification:** Developed a document type classification model using **SwinV2 (ViT)** architecture, automating the identification process with high accuracy and reducing manual review times.
- **Content Extraction:** Deployed **LLM** and **OCR**-based extraction techniques to capture critical document data, improving accuracy and reliability in data extraction workflows.

Technical Skills:

- **Machine Learning & LLMs:** Proficient with OpenAI and Hugging Face Models
- **OCR Technology:** Experienced with Google OCR and Lens
- **Cloud Computing:** Skilled in AWS (Lambda, S3, CloudWatch) for scalable deployment
- **API Development:** Competent in FastAPI and Django for robust API creation
- **Cross-functional Collaboration:** Effectively partnered with developers and data validators to ensure comprehensive project delivery

---

Data Scientist

La Poste Groupe | September 2021 – February 2024 | Nantes, France (Hybrid)

Applied **machine learning**, **NLP**, and **graph databases** in **network security** and **financial fraud detection**.

Key Contributions:

- **Network Activity Forecasting:** Identified patterns and predicted anomalies to enhance security measures using **Temporal Fusion Transformer (TFT)**.
- **Graph Database Utilisation:** Leveraged **Neo4j** to map abnormal banking activities, improving fraud detection.
- **Complex Relationship Analysis:** Analysed and visualised relationships within banking networks to understand illicit activity patterns.
- **NLP in Cybersecurity:** Applied **NLP** techniques to command-line interfaces for threat detection, focusing on **Juniper commands**.

- **Deep Learning Models:** Developed **LSTM** and **Autoencoder** models using **PyTorch** for network security analysis.
- **Log Data Analysis:** Used **Splunk** for collecting and analysing network device logs, crucial for incident detection.

Technical Skills:

- **Programming Languages:** Python
- **Deep Learning & NLP:** PyTorch, LSTM, Autoencoders
- **Graph Databases:** Neo4j
- **Cybersecurity Tools:** Splunk
- **Operating Systems:** Linux (CentOS)

Data Scientist Intern

Institut de Neurosciences de la Timone (CNRS) | February 2021 – May 2021 | Marseille, France

Gained experience in **computer vision** and **deep learning** through hands-on projects.

Key Projects:

- **Inhibition of Return Mechanism:** Implemented using **PyTorch** to build advanced computer vision models.
- **Dataset Modification:** Enhanced the **MNIST** dataset with **NumPy** to simulate background motion.
- **Experimental Design:** Conducted human-equivalent experiments using **PsychoPy** to assess model performance.

Technical Skills:

- **Deep Learning:** PyTorch
- **Data Manipulation:** NumPy
- **Experimental Tools:** PsychoPy

Lab Assistant

University of Newcastle | February 2018 – September 2018 | Central Coast - Ourimbah (On-site)

Prepared participants for **EEG** experiments and analysed resulting data.

Responsibilities:

- **Participant Management:** Screened and recruited participants for EEG studies.
- **EEG Preparation:** Fitted electrodes and prepared participants for recordings.
- **Psychological Assessments:** Administered and scored evaluations of anxiety, depression, and psychopathy.

- **Data Analysis:** Used **Principal Component Analysis (PCA)** to identify patterns in EEG data.

Technical Skills:

- **Data Analysis:** PCA
- **Neuroscience Tools:** EEG data collection and preparation

---

Education

Deakin University

Master's Degree in Data Science | February 2019 – October 2021

- Mastered concepts of **data acquisition**, **storage**, **processing**, and **analysis**.
- Gained proficiency in **Python** and **R** for manipulating and analysing large datasets.
- Applied machine learning algorithms: **regression**, **classification**, **clustering**, and **deep learning**.
- Understood ethical and legal considerations related to **data privacy** and **security**.
- Designed and implemented **data-driven solutions** to complex business problems.
- Communicated technical concepts to both technical and non-technical stakeholders.
- Utilised **data visualisation** and **storytelling** to effectively communicate insights.
- Collaborated with professionals to drive **data-driven decision-making**.

Skills Acquired:

- **Programming Languages:** Python & R
- **Data Visualisation:** Tableau
- **Machine Learning Frameworks:** TensorFlow & Pytorch

Aix-Marseille University

Master's Degree in Cognitive Science | 2019 – 2021

- Studied fundamental concepts, theories, and methods of **cognitive science**.
- Applied advanced statistical and computational methods to analyse complex datasets.
- Conducted experimental research: **behavioural experiments**, **neuroimaging**, **eye-tracking**.
- Understood neural and computational mechanisms underlying cognitive processes.
- Analysed data from brain imaging techniques: **fMRI**, **EEG**, **MEG**.

- Designed **cognitive models** to simulate and predict human behaviour and cognition.

Grade: Distinction

Skills Acquired:

- **Programming Languages:** Python, R
- **Machine Learning Frameworks:** TensorFlow

University of Newcastle

Bachelor of Psychological Science | February 2016 – November 2018

- Specialised in **psychological science** with a strong foundation in **research** and **statistics**.
- Completed a rigorous 240-unit program emphasising **research methods** and **cognitive psychology**.
- Earned a place on the Faculty of Science and IT **Commendation List** in 2016.
- Achieved high distinctions in **Advanced Research Methods**, **Statistics**, and **Cognitive Psychology**.

Grade: Distinction

Skills Acquired:

- **Statistical Analysis:** Advanced research methods
- **Psychology Expertise:** Cognitive psychology, social and organisational psychology

---

Skills

- **Programming Languages:** Python, R
- **Machine Learning & AI:** OpenAI, HuggingFace, PyTorch, TensorFlow, LSTM, Autoencoders
- **Data Extraction & OCR:** Google OCR
- **Data Visualisation:** Tableau
- **Cloud Computing:** AWS Lambda, S3
- **Databases:** PostgreSQL, Neo4j
- **API Development:** FastAPI
- **Cybersecurity Tools:** Splunk
- **Operating Systems:** Linux (CentOS)
- **Data Analysis Tools:** NumPy, PsychoPy
- **Research & Statistics:** PCA, experimental design
- **Collaboration & Communication:** Cross-functional teamwork, data validation, project management

---