# 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:**

- **Large Language Models & Vision Models:** Expertise in OpenAI and Hugging Face models for natural language processing and visual recognition tasks, with strong knowledge of **SwinV2** and ViT architectures.
- **OCR & Computer Vision Tools:** Proficient with **Google OCR** and **Google Lens** for document digitisation and data extraction.
- **Cloud Infrastructure:** Skilled in **AWS** (Lambda, S3, CloudWatch) for scalable, cost-effective deployment and monitoring of machine learning pipelines.
- **API Development:** Proficient in **FastAPI** and **Django** for building and managing robust, high-performance APIs.
- **Cross-Functional Collaboration:** Partnered closely with developers, data validators, and stakeholders to ensure seamless project execution and delivery.

---

## Data Scientist – Network Security & Fraud Detection

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

Specialised in machine learning, NLP, and graph analytics to bolster network security and advance financial fraud detection initiatives.

**Key Contributions:**

- **Network Activity Forecasting:** Implemented the **Temporal Fusion Transformer (TFT)** in **PyTorch** to predict network anomalies, significantly improving the accuracy of preemptive security measures.
- **Graph-Based Fraud Detection:** Utilised **Neo4j** to construct and query graph-based models of banking transactions, enhancing the detection of complex fraud patterns by uncovering hidden relationships.

- **Banking Network Analysis:** Designed and visualised intricate networks of financial transactions, helping to trace and understand behavioral patterns associated with fraudulent activities.
- **NLP for Threat Detection:** Applied advanced **NLP** techniques to parse and analyse **Juniper command-line logs**, enabling automated detection of suspicious command patterns in cybersecurity contexts.
- **Deep Learning for Anomaly Detection:** Built and fine-tuned **LSTM** and **Autoencoder** models for anomaly detection in network security, leveraging **PyTorch** for high-performance training and inference.
- **Log Data Analysis & Incident Response:** Streamlined log analysis workflows with **Splunk**, creating automated alerts and dashboards for continuous monitoring and faster incident detection.

**Technical Skills:**

- **Programming Languages:** Proficient in **Python** for data science and model development.
- **Machine Learning & Deep Learning Frameworks:** PyTorch, LSTM, Autoencoders, TFT for time-series and anomaly detection.
- **Graph Databases:** Expertise in **Neo4j** for fraud detection and relationship analysis.
- **Cybersecurity & Monitoring Tools: Splunk** for log management, analytics, and incident detection.
- **Operating Systems:** Skilled in **Linux (CentOS)** for secure and efficient server management.

---

## Data Scientist Intern – Computer Vision & Deep Learning

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

Developed and deployed computer vision models, gaining hands-on experience in deep learning and experimental design for neuroscience applications.

**Key Projects:**

- **Inhibition of Return Mechanism Simulation:** Designed and implemented neural models in **PyTorch** to replicate and study visual attention mechanisms, such as the inhibition of return in image sequences.
- **Dataset Augmentation & Simulation:** Applied **NumPy** to modify the **MNIST** dataset, simulating background motion to enhance model robustness and evaluate performance in dynamic environments.
- **Experimental Design & Human-Equivalent Testing:** Conducted controlled experiments using **PsychoPy** to assess neural network performance against human visual attention benchmarks.

**Technical Skills:**

- **Deep Learning Frameworks:** PyTorch for model design, training, and evaluation.

- **Data Manipulation & Augmentation:** Advanced **NumPy** operations for dataset modifications and synthetic data generation.
- **Experimental & Simulation Tools: PsychoPy** for designing and running cognitive and behavioral experiments to validate model efficacy.

---

### Lab Assistant – EEG Data Analysis

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

Conducted EEG experiments, preparing participants and analysing neural data for research studies.

**Responsibilities:**

- **Participant Management:** Screened, recruited, and onboarded participants for EEG studies, ensuring eligibility and compliance with study protocols.
- **EEG Setup & Monitoring:** Fitted electrodes, optimised signal quality, and prepared participants for EEG recordings to ensure accurate data capture.
- **Psychological Assessments:** Interpreted assessments for anxiety, depression, and psychopathy, contributing to data-driven insights.
- **Data Analysis:** Applied **Principal Component Analysis (PCA)** and other statistical methods to extract meaningful patterns from EEG data, enhancing data interpretation and research outcomes.

**Technical Skills:**

- **Data Analysis & Statistics:** Principal Component Analysis (PCA), signal processing techniques for EEG data.
- **Neuroscience & Data Acquisition:** Hands-on EEG setup and preparation, electrode placement, signal optimisation.
- **Psychometric Testing & Evaluation:** Proficient in administering and scoring standardised psychological assessments.

---

## Education

### Deakin University

Master of Data Science

**Key Skills Acquired:**

- **Data Analytics & Machine Learning**: Mastery in machine learning, statistical data analysis, and real-world analytics through coursework in Machine Learning, Statistical Data Analysis, and Mathematics for AI. Skilled in practical data wrangling for structured and unstructured data.

- **Software Development & IT Foundations**: Proficient in object-oriented programming, database fundamentals, and software requirements analysis, supported by courses such as Object-Oriented Development, Database Fundamentals, and Web Technologies.
- **Advanced AI Techniques**: Extensive experience in Bayesian learning, deep learning, and modern data science, acquired through Bayesian Learning, Deep Learning, and Modern Data Science modules.
- **Collaborative Project Execution**: Strong project management and teamwork skills gained through team-based projects, including Project Management and Practices and Execution and Delivery, focused on managing data science projects from planning through to delivery.

---

## OpenClassrooms

Machine Learning Engineer Graduate Diploma (Master's Level, Bac +5)

**Key Skills Acquired:**

- **Data Preparation & Analysis**: Cleaning, transforming, and visualising structured and unstructured data.
- **Predictive Modeling & Machine Learning**: Designing and deploying predictive algorithms, including supervised and unsupervised models, deep learning (CNNs for image classification), and NLP for text processing.
- **MLOps & Deployment**: End-to-end lifecycle management of models using cloud platforms (AWS, Google Cloud), version control (Git), and monitoring in production.
- **Advanced Technical Proficiencies**: Python (Jupyter, Git, Streamlit), MLFlow, Databricks, Hadoop/PySpark, and cloud computing for scalable data solutions.
- **Project-Based Experience**: Completed industry-aligned projects such as energy consumption forecasting, e-commerce client segmentation, and predictive modeling of user behaviors.

---

## OpenClassrooms

Data Analyst Graduate Diploma (Bachelor's Level, Bac +3/4)

**Key Skills Acquired:**

- **Data Collection & Analysis**: Aggregating, cleaning, and interpreting data to derive insights.
- **Statistical Analysis & Hypothesis Testing**: Conducting descriptive statistics, statistical modeling, and testing hypotheses for data-driven decisions.
- **Data Visualisation & Reporting**: Creating impactful reports, dashboards, and presentations for various stakeholders.
- **Technical Proficiencies**: SQL, Python, R, data visualisation tools (e.g., Tableau), and statistical libraries.

- **Project-Based Learning**: Completed multiple hands-on projects, including public health studies, market analysis, sales data analysis, and predictive modeling for business insights.

---

## University of Newcastle

Bachelor of Psychological Science

**Key Skills Acquired:**

- **Research Methods & Statistics**: Developed expertise in research design, data analysis, and psychological measurement through courses like Advanced Research Methods and Statistics and Advanced Psychological Measurement.
- **Core Psychological Knowledge**: Gained in-depth knowledge in cognitive psychology, biological psychology, and mental health through core courses, including Cognitive Psychology, Biological Psychology, and Introduction to Mental Health.
- **Advanced Specialisations**: Specialised in applied areas like health psychology, social and organisational psychology, and developmental psychology, covered in Advanced Health Psychology and Advanced Developmental Psychology.
- **Applied Learning and Professional Practice**: Completed Work Integrated Learning in Psychology, providing hands-on experience in real-world settings to enhance employability and practical skills in psychology.

---