Manuel Arnol Fokam

COMPUTER SCIENCE M.SC CANDIDATE

arnolfokam23@gmail.com @arnolfokam
@ArnolFokam arnolfokam.github.io

SKILLS & TOOLS

Natural Language Processing | Algorithms & Data Structures | Scikit-Learn | PyTorch | Self-Supervised Learning (computer vision) | Tensorflow | JavaScript | Pandas

ACADEMIC EXPERIENCE

University of Buea, Cameroon. 2016 - 2019

B.Sc. Applied Physics (Minor Computer Science), GPA: 3.12/4.00

- **Activities:**
  - Academic Officer, University of Buea Association for the Advancement of Physics
  - Course delegate: CSC208 (Programming in C and Python), PHY302 (Analogue Electronics Laboratory)
  - Team Member, Applied Physics Basketball Team (B-Sera Basketball Competition)
- **Final Year Project:** Introduction to Quantum Computing (Theory and Applications)

University of the Witwatersrand, Johannesburg. 2021 - 2021

B.Sc. (Hons) Computer Science, (cum laude)

- **Research Topic:** Influence of contrastive Learning on Source Code Plagiarism Detection

University of the Witwatersrand, Johannesburg. 2022 - Present

M.Sc. (Dissertation) Computer Science

- **Research Topic:** Adversarial Gradient-based Data Augmentation for Self-Supervised Learning
- **Activities:**
  - Organizer of the **RL Workshop** at the **Deep Learning Indaba 2022**
  - Member of the **Procedural Content Generation** Interest Group

PROFESSIONAL EXPERIENCE

Software Engineer Intern (Remote), Boostasoft (Startup), France (Feb, 2020 – Feb, 2021)

- Evolved in an autonomous environment where I learned to translate the stakeholders requirements into product roadmaps. In addition, I learned to evolve these roadmaps into deployable MVPs.
- **Technology Used:** Java (Spring Boot), JavaScript (React, React Native)
- **Learned Skills:**
  - Project Management
  - Working in a Startup culture that favours autonomy.
- **Products built:**
  - GetJobs Platform (https://www.getjobs.careers/),
  - CovidAlert App (https://boostasoft.com/covid-alert/index.html)

PUBLICATIONS

Influence of Contrastive Learning on Source Code Plagiarism Detection through Recursive Neural Networks. *Manuel A. Fokam, Ritesh Adoojha*

PROJECTS

DNA Gate

- Built a SaaS platform that permits the rapid integration of biometric authentication mechanisms (face and voice) through API calls.
- **Technology Used:** Python (FastAPI, NumPy, OpenCV), JavaScript (React, MongoDB, ExpressJS)
- **Highlights:**
  - Built a module to handle API keys generation and track usage history.
- **Key Takeaways:**
  - Deployed ML solutions can be enhanced by pre-processing inputs before inference:
    - An eyes detector or lighting enhancement algorithm can be used as preprocessing steps before facial recognition.
    - A noise reduction algorithm can be used as a preprocessing step before voice recognition.

COVID-19 infection segmentation from Pre-Trained encoders

- Built a machine learning project that evaluates the performance of U-Net with various pre-trained encoders on a medical image segmentation task. The task is to identify COVID-19 infections from lungs' CT scans.
- **Technology Used:** Python (Pandas, NumPy, OpenCV, Jupyter Notebook)
- **Highlights:**
  - Made a class that handles the retrieval and preprocessing of CT scans from the **Cancer Imagine Archive** (https://www.cancerimagingarchive.net/) API.
- **Key Takeaways:**
  - CT scans need more advanced preprocessing techniques because:
    - Their pixel range known as Hounsfield units (HU) is larger than RGB images.
    - They are inherently volumetric in nature to contain the full scan of a lung or other body parts.

Battlefield Honor

- Built a multiplayer battle royale first-person 2D shooter game.
- The project was made by a team of 3 members, each focusing on a specific component of the game:
  - Game Server (my part)
  - Game Client
  - Maps of the Game
- **Technology Used:** JavaScript (ColyseusJS - server, PhaserJS - client, ReactJS - landing page)
- **Highlights:**
  - Dealt with latency by giving different responsibilities to client and server.
    - Client: submit actions to server and listen for game state changes.
    - Server: Listen to action from client and update game state accordingly.
- **Key Takeaways:**
  - The success of real-time multiplayer games highly depends on how you deal with latency when designing server/client communications.
  - An effective is a team where each member have non-overlapping skills and are the best at what they do.

NER Annotations Density vs Pre-trained Language Models

- Built Machine Learning project that analysis the performance of Language models on Named Entity Recognition(NER) Tasks when the quality of the dataset is altered.
- The quality of the dataset here represents the minimum amount of misclassification allowed per sentences.
- **Technology Used:** Python (Hugging Face, Jupyter Notebook)
- **Highlights:**
  - Built a function that could create a subset of a NER dataset with variable quality.
  - Presented a poster of this work at the **Deep Learning Indaba 2022**.
- **Key Takeaways:**
  - Open-Sourced work is a great way to get feedbacks for future improvements.
  - Model cards are very efficient ways to share ML research.

Source Code Plagiarism Detector

- Built a source code plagiarism detector that uses contrastive learning to identify similar code snippets.
- The model used to get the latent representations of the code snippets is a Tree-based Neural Network.
- The main challenge was to port the existing open-sourced code from PyTorch to TensorFlow.
- **Technology Used:** Python (TensorFlow, Pandas)
- **Highlights:**
  - Got the paper for this project accepted at an African IT conference.
- **Key Takeaways:**
  - We can leverage the abstract syntax trees of code snippets for representation learning in programming languages.

AWARDS

DeepMind Scholarship - University of the Witwatersrand. 2022

Dean's List Honours - University of the Witwatersrand. 2021

Postgraduate Merit Award- University of the Witwatersrand. 2021

1st Prize InnovaTech Hackathon - Yaounde, Cameroon. 2020

ON-GOING INITIATIVES

Creation of a Tech Interview Prep Interest Group

- Establishing an interest group to help students prepare for tech internships at big technological companies.
- **Reason:** We need more representation of people coming from developing countries at these companies.

Creation of an AI Competition Interest Group

- Establishing an interest group to set up student form teams to solve real-world AI problems advertised on different AI completion websites.
- **Reason:** This is an excellent way to gain experience with AI tools and their application in real-world scenarios.

Creation of Paper Reading Club

- Establishing a paper reading club at the Computer Science Department of my current University.
- **Reason:** This is an excellent way to stay up-to-date with the current state of AI/ML research,

HOBBIES & INTERESTS

Blogging | BasketBall | Ancient Antiquity (casual)

SPEAKS

English (fluent) | French (fluent)