Ben KABONGO

Student, Paris, France
kabongo.ben025@gmail.com — +33 7 58 52 58 85 — LinkedIn — Github

RESEARCH INTERESTS

Artificial Intelligence, Deep Learning, Machine Learning, Natural Language Processing, Recommender Systems

EDUCATION

Ecole Normale Supérieure Paris-Saclay, Orsay, France 2023 — 2024
Master's degree in M2 Mathematics Vision Learning (MVA)

Sorbonne University, Paris, France 2022 — 2023
Master's degree in M1 Data Learning Knowledge (DAC)

Paris-Saclay University, Orsay, France 2022 — 2023
Master's degree in M1 Data Science

Sorbonne University, Paris, France 2021 — 2022
Bachelor's degree in L3 Computer Science

University of Caen Normandy, Caen, France 2019 — 2021
Bachelor's degree in L1-L2 Computer Science

AWARDS

Google DeepMind Scholarship September 2023
Google DeepMind scholarship obtained for my year of the Mathematics Vision Learning (MVA) master's degree at Ecole Normale Supérieure Paris-Saclay.

SaclAI-School Scolarship September 2022
Scholarships for undergraduate and graduate students in the Artificial Intelligence and Data Science programs at Université Paris-Saclay.

INTERSHIPS

AgroParisTech (MIA Paris-Saclay) & Onepoint (TALia) Paris, France
Recommender Systems and Text Data Analysis April 2024 — October 2024
The aim of this project is to work at the interface between recommender systems and natural language processing, in order to exploit text data to propose new, efficient and explainable recommendation models.

- **Supervision**: Vincent Guigue (MIA Paris-Saclay, AgroParisTech) and Pirmin Lemberger (Onepoint)
- **Keys concepts**: Recommender Systems, Natural Language Processing, Large Language Models, Sentiment Analysis/Opinion Mining, Explainability
- **Programming/Software**: Python, Pytorch, Scikit-suprise

Machine Learning and Deep Learning for Information Access (MLIA) Paris, France
Personalized data-to-text neural generation June 2023 — August 2023
The aim of the internship is to develop a neural data-to-text system capable of customizing text generation.

- **Supervision**: Laure Soulier (ISIR/MLIA, Sorbonne University), Antoine Gourou (Université de Saint Etienne) and Christophe Gravier (Télécom Saint-Etienne)
- **Keys concepts**: Data-to-text, Personalization, Natural Language Processing
- **Programming/Software**: Python, Pytorch

SKILLS

- **Programming**: Python, Java, JavaScript, C, C++, HTML, CSS, SQL
- **Software**: PyTorch, Sklearn, Numpy, Pandas, Matplotlib, Scipy, Seaborn, React, NodeJS, JDBC

PROJECTS

ALTeGraD Challenge Molecule Retrieval with Natural Language Queries January 2024 — February 2024
Mathematics Vision Learning, Ecole Normale Supérieur Paris-Saclay
The goal of this project is to study and apply machine learning/artificial intelligence techniques to retrieve molecules (graphs) using natural language queries.

- **Key concepts**: Natural Language Processing, Graphs
- **Programming/Software**: Python, Pytorch
- Kaggle, Github

Decision Transformer: Reinforcement Learning via Sequence Modeling December 2023 — January 2024
Mathematics Vision Learning, Ecole Normale Supérieur Paris-Saclay
As part of the Deep Learning course I took at the MVA master's program, we did a project on a research paper. I worked on the "Decision Transformer: Reinforcement Learning via Sequence Modeling" paper [L. Chen and al.].

- **Key concepts**: Reinforcement learning, Natural Language Processing, Transformer.
- **Programming/Software**: Python, Pytorch, Numpy
- Paper, Github

MAGMA November 2023 — December 2023
Mathematics Vision Learning, Ecole Normale Supérieur Paris-Saclay
As part of the Probabilistic Graphical Models course I took at the MVA master's program, groups of 3 students worked on the "MAGMA: inference and prediction using multi-task Gaussian processes with common mean" paper [Arthur Leroy, Pierre Latouche, Benjamin Guedj, Servane Gey].

- **Key concepts**: Multi-task learning, Gaussian Processes, EM algorithm, Common mean process, Functional data analysis.
- **Programming/Software**: Python, Numpy, Sklearn, Scipy, Matplotlib
- Paper, Github

DILATE November 2023 — December 2023
Mathematics Vision Learning, Ecole Normale Supérieur Paris-Saclay
As part of the Machine Learning for Times Series course I took at the MVA master's program, groups of 2 students worked on the "Shape and Time Distortion Loss for Training Deep Time Series Forecasting Models" paper [Vincent Le Guen, Nicolas Thome].

- **Key concepts**: Time series forecasting
- **Programming/Software**: Python, Numpy, PyTorch
- Paper, Github

PLDAC : Are language models able to generate instructions for robots? January 2023 — May 2023
Data Learning Knowledge, Sorbonne University
The aim of this project is to study the ability of large language models to generate coherent instructions for robots, by determining whether they have the necessary semantic knowledge.

- **Supervision**: Laure Soulier and Nicolas Thome
- **Key concepts**: Natural language processing, Reinforcement Learning
- **Programming/Software**: Python, Numpy, Sklearn, Scipy, PyTorch

Projet Réseau de neurones - DIY January 2023 — May 2023
Data Learning Knowledge, Sorbonne University
The aim of this project was to implement a neural network. The implementation is inspired by older versions of pytorch (in Lua, before autograd) and similar implementations, which allow for highly modular generic networks.

- **Key concepts**: Machine Learning, Deep Learning
- **Programming/Software**: Python, Numpy, Matplotlib
- Github

Combining deep reinforcement learning and evolutionary methods January 2022 — May 2022
L3 Computer Science, Sorbonne University
CEM-RL and ERL are two algorithms that combine reinforcement learning and evolutionary methods. The aim of this project is to implement a new variant of these algorithms that incorporates the best elements of both combinations, and then to evaluate this variant.

- **Supervision**: Olivier Sigaud
- **Key concepts**: Reinforcement learning, Evolutionary methods
- **Programming/Software**: Python, PyTorch, Stablebaslines3, Gym, Salina
- Github

Creation of a website like Twitter January 2022 — May 2022

L3 Computer Science, Sorbonne University

As part of my L3 Computer Science course, we were asked to design a Twitter-like website.

- **Programming/Software**: HTML, CSS, JavaScript, React, NodeJS, SQL, NoSQL, JSON
- Github

3D rendering using ray tracing January 2021 — May 2021

L2 Computer Science, University of Caen

CEM-RL and ERL are two algorithms that combine reinforcement learning and evolutionary methods. The aim of this project is to implement a new variant of these algorithms that incorporates the best elements of both combinations, and then to evaluate this variant.

- **Key concepts**: Computer Vision, Ray tracing
- **Programming/Software**: Java, JavaFx
- Github