Baptiste Moreau-Pernet

Data Scientist ▪ Learning Engineer

(412) 626 1719
baptiste.moreaupernet@gmail.com
baptistemp.github.io

ACADEMIC EXPERIENCE

Natural Language Processing Course   *CMU | Spring 2021*
- Implemented word edit distances for text correction, N-grams and Bayesian text classifiers, Viterbi algorithm for Part-of-Speech tagging.
- Building a question generation/answering model in a final group project.

Educational Goals, Instruction & Assessment Project   *CMU | Fall 2020*
- Designed a 10-hour course for 2nd graders for algebraic thinking, following backward design methodology.
- Elaborated an evaluation research module for course improvement.

Personalized Online Learning Project   *CMU | Fall 2020*
- Designed an adaptive educational course in Geography for college students using Cognitive Task Analysis, improved it with A/B testing.
- Implemented the course in a digital cognitive tutor in Javascript (nools)

Capstone Project   *CentraleSupélec | 2017 to 2019*
- Co-designed a skill-based job description ontology to match job offers to job seekers' de-identified resumes in a 2-year, 4-person project.
- Predicted job turnover with Machine Learning classifiers using Python, sklearn and pandas, which resulted in a 0.98 accuracy.

Innovation Project   *CentraleSupélec | Spring 2019*
- Co-designed a neural network-driven autonomous car using Raspberry Pi 3 and Arduino in a 2-person semester-long project.
- Developed a convolutional Neural Network model with TensorFlow to drive the car using floor pictures, which achieved a 0.91 accuracy.

PROFESSIONAL EXPERIENCE

Research Assistant at Carnegie Mellon University, Pittsburgh, PA
Fall 2020 to Spring 2021
- Extracted valuable insights from a SQL database storing chatbot conversational metadata to refine interaction design using Numpy, in a research project to train low-resourced Ivoirian teachers.
- Created data visualizations to monitor the deployment for 500 teachers.

Data Scientist Intern at Calyps S.A, Sion, Switzerland
March 2020 to July 2020
- Designed and implemented a patients' hospital stay software to improve codification efficiency and save time for clinical professional coders.
- Trained a BERT Deep Learning classifier on 30K+ natural language medical notes with 400+ target classes, using state-of-the-art Natural Language Processing on PyTorch, and achieved a 60% accuracy.
- Ran the software's operational implementation for a private clinic.

Research Intern at Cognitive Science & Psycholinguistics Lab, ULM, Paris
September 2019 to January 2020
- Developed a R statistical matching model to identify grade repetition's effects, and ran the study on a 30K+ Junior High School student cohort.
- Concluded on the negative average effect of grade repetition on learning.

EDUCATION

Carnegie Mellon University
School of Computer Science
Master of Educational Technology & Applied Learning Science
August 2021

CentraleSupélec
Master of Science & Engineering
June 2020

CentraleSupélec
Bachelor of Science & Engineering
June 2018

SKILLS | Data Science
Languages
Python, JavaScript, SQL, R
Algorithms
Deep Learning, Natural Language Processing
Libraries
Pandas, Numpy, seaborn, Keras, PyTorch, TensorFlow

SKILLS | Learning Engineer
Course design
cognitive task analysis, A/B testing, backward design
Online learning
cognitive tutors

HONORS
Rotary International Global Grant Scholarship
Rotary's Development Axe of Education
CMU | 2020 to 2021

OTHER INTERESTS
Multicultural experiences
Competitive rowing
Black belt in Karate
Ski at a competitive level
Guitar player