Clement Mugenzi

Email: cm3928@cumc.columbia.edu   **Portfolio**: Website
Cell Phone: (347) 703 2776   **LinkedIn:** Profile

EDUCATION

Columbia University, New York, NY - Master of Science, Biostatistics   Expected: May 2021
Relevant Coursework: Data Science, Statistical Methods, Probability, Statistical Inference
University of Central Arkansas, Conway, AR **-** Bachelor of Science**,** Chemistry   2016

SKILLS

Programming: Python, R, SAS, LaTeX
Databases: MySQL, PostgreSQL
Libraries & Packages: Pandas, Numpy, Matplotlib, Scikit-learn, Tensorflow (keras), Tidyverse, Caret
Languages: English (Native), French (Native), Kinyarwanda (Native), Spanish (Basic)

PERSONAL PROJECTS

1. **Green Taxi: Predicting Percent Tip**
Machine learning tools in Python were used to predict percentage tip a driver would expect on each trip. Several features both categorical and continuous were considered as independent predictors such as total amount payed, trip distance, payment type, speed, etc. This project followed four main sections: Data Cleaning, Feature Engineering, Exploratory Data Analysis, and Model Building. A Gradient Boosting classification model was optimized (with a 96.1% AUC and 96.6% Accuracy) to predict whether or not a tip was provided, followed by a regression Random Forest model which then estimated the percentage tip given the tip was provided (MSE was 0.8).
See sample code and analysis here: Green Taxi Project

2. **Titanic: Survival Analysis**
Machine learning tools in R studio were used to predict individuals more likely to survive the crash. Several features were considered as independent predictors such as social class, gender, age, marital status, etc. The data was acquired from the original dataset and variables were cleaned and validated for an easy and comprehensive exploratory data analysis. A classification model with a AUC of 87.1% was finally built by optimizing a random forest algorithm.
See sample code and analysis here: Titanic Project

WORK EXPERIENCE

Allure Labs, San Francisco, CA   09/2018 – 05/2019
   QC Chemist
- Responsible for providing support to any investigation regarding Lab deviation or Non-conformance, Out of specification and change control procedure.
- Worked closely with the research and development team to provide product specifications for newly designed cosmetics.
- Operated both HPLC and ICP instruments to test for the presence of Zinc and Titanium oxides in OTC products.

American Interplex Corp., Little Rock, AR   04/2017 – 11/2017
Analytical Chemist
- Used ChemStation software to operate Agilent GC-MS, HPLC, FIDs instruments and to analyze samples then report data to clients in accordance to guidelines set forth by the EPA.
- Analyzed for Herbicides, Pesticides, PCBs, and amino acids using both GC-MS and HPLC.
- Built calibration curves (to construct analytical methods), using regression analysis, from which the instruments would infer and quantitate unknown samples.
- Devised a model to detect and/or drop outliers and together with a proper maintenance of the instrument contributed to reducing the processing time by 50%.

University of Central Arkansas, Conway, AR 01/2013 – 04/2016
    Research Assistant

- Conducted chemistry research on the construction of dynamic coordination polymer materials where one to two dimensional polymers were built.
- Presented my research findings at the ACS meeting in San Diego (March 2016)

PUBLICATION

Synthesis and characterization of divalent metal complexes with bipyridylamide ligands, Clement Mugenzi *et al*, Journal of Coordination Chemistry 2015.

- Synthesized 2D and 3D polymers and used CrystalMaker to analyze their structure.
- Using CrystalMaker as repository, I surveyed the literature to compare our discoveries with previous researches in order to better understand our polymers' crystal structure.
- CrystalMaker is a visualization tool that builds, displays, and helps manipulate all kinds of crystal structures.
- **Transferable skills:** Computation, Critical Reasoning, Research Design, Data Visualization.
- **Reference:** Research Paper

PROFESSIONAL ORGANIZATION

- Member: Health Analytics Club - Columbia University   2019 – Present
- Member: Computer Club – Biostatistics Department   2019 - Present
- Member: American Chemical Society   2013 - 2016
- Member: Inspire Scholars Foundation   01/2014 - 07/2014

AWARDS & HONORS

- Taub Institute Award at Columbia   2019
- Rwanda Presidential Scholar   2012-2016
- The Nicole Wable Hatfield scholarship.   2014