Michael Armanious

https://www.michaelarman.github.io
michael.arman7@gmail.com | +1 647 997 3306 | linkedin.com/in/michael-armanious

SKILLS

PROGRAMMING

Python and iPython Notebook
R and R Studio
SQL and CQL
SAS
HTML and Markdown

MACHINE LEARNING

Regression
Decision Trees
Support Vector Machines
Random Forests & XGBoost
Cluster Analysis
Neural Networks/Deep Learning:
Image Classification
Text Classification
Audio Classification

TOOLS & SOFTWARE

SQL (PostgreSQL, MySQL)
IBM Netezza
SAS & Neo4j
Git and Github
Jupyter, Google-Colab
Anaconda + packages
ggplot2, matplotlib, seaborn
WEKA

EDUCATION

MSC STATISTICS

Carleton University
Sept 2019 | Ottawa, ON Master's Project: Men's U-Sports Basketball Analysis

BSC, MATH & STATS

Carleton University
May 2018 | Ottawa, ON

CERTIFICATES

Deep Learning, Tensorflow, Computer Vision, NLP
Coursera, DataCamp, Udacity

LINKS

Github:// michaelarman
LinkedIn:// michael-armanious

EXPERIENCE

CANADA REVENUE AGENCY | Data Scientist

May 2017 – Present | Ottawa, ON

- Tools: SAS, Anaconda, R, Python; numpy, pandas, sci-kit learn, tensorflow, keras, pytorch, Neo4J, Netezza SQL, Excel, HTML
- Collected, studied, and interpreted large datasets of taxpayer information and generating Relational Graph Databases using Neo4j and SAS to classify and predict compliance.
- Creating classification and prediction models on large imbalanced datasets using various machine learning algorithms such as Logistic Regression, Random Forests, XGBoost, Neural Networks, etc. to classify and predict which taxpayers should be audited.
- Conducted reports to management about graph databases, compliance studies and risk assessment.

CARLETON UNIVERSITY | Research & Teaching Assistant

Sept 2018 – Sept 2019 | Ottawa, ON

- Nominated for *Outstanding TA Award*
- Data Mining - STAT 4601/STAT 5703: Tasked with assisting students with Data Mining projects using R Studio and teaching them proper data mining techniques including data cleaning, data preparation, data analysis and write-ups.
- ANOVA and Experimental Design: STAT 3504 & Sampling Methodology: STAT 3507: Taught tutorials and demonstrated how to use SAS in order for the students to complete their assignments
- Introduction to Statistical Modelling: STAT 2507: Taught tutorials for students taking intro to statistics and showed them how to use Minitab and how to interpret findings
- Implemented several data mining algorithms in datasets such as OUA (Ontario University Athletics) basketball using the Synergy Sports Technology to.create reports to the coach of the Carleton Basketball team
- Scraped and cleaned raw data from the OUA website and Synergy database to use for analysis. Incorporated classification and clustering algorithms in order to gain meaningful insights on Basketball data.
- Explored data to search for structure/patterns/factors/clusters by performing visualization and summary techniques.

RESEARCH & PROJECTS

- Tools: R , R Markdown, RShiny, SQL, Python; BeautifulSoup, scrapy, selenium, requests, scikit-learn, pandas, numpy, matplotlib, tensorflow, pytorch, bokeh, librosa, OpenCV
- Master's Project: Data Mining on OUA Basketball
- Project: Audio Classification of Accents using Deep Learning
- Project: Steganalysis
- Project: Heart Disease Analysis

Relevant Coursework

Categorical Data Analysis, Bayesian Analysis, High Dimensional Analysis, Advanced Data Mining, Time Series Forecasting, Multivariate Analysis, Survey Sampling, Combinatorial Optimization, Regression Analysis, Stochastic Processes & Applications, Analysis of Variance & Experimental Design