Justin **Charbonneau**

2-542 Montée Paiement, Gatineau, QC, J8R 0G8

819-319-3622 | justincharbonneau.ai@gmail.com | justincharbonneau.github.io | in/charbonneau217/

Work Experience

Research Assistant | Knowledge Discovery and Data Mining (KDD) Lab *Ottawa, Ontario*
UNIVERSITY OF OTTAWA *Sep. 2019 - PRESENT*
- Directed the data science project *Text Mining for Intelligent Market Trading*
- Research on machine learning algorithms for stock market price movements
- Helped lab members with Python

Data Scientist (CO-OP) *Ottawa, Ontario*
ARIO PLATFORM *Jan. 2019 - Aug. 2019*
- Created an R package for collecting Google Trends data and aggregating results across time and location which is ready to be used
- Performed data collection, data exploration and data cleaning on a regular basis in R
- Developed an image processing and computer vision pipeline to extract information from image files using Python and open-source libraries such as OpenCV for its image processing capabilities; applied and compared several OCR engines (AWS, Google and Microsoft)
- Demonstrated a Jupyter Notebook to the staff to showcase the image processing pipeline
- Documented all work in Confluence and generated reports using R Markdown and R Flex Dashboards

Programmer *Ottawa, Ontario*
FISHERIES AND OCEANS CANADA *Jan. 2018 - Dec. 2018*
- Improved the accuracy of the information stored in our database by developing a PHP web application that identified discrepancies and allowed users to validate the information and track the validation progress
- Exploited APIs to update the information in the configuration management database
- Displayed summary data using Google Chart API

Junior System Analyst *Ottawa, Ontario*
FISHERIES AND OCEANS CANADA *Apr. 2017 - Dec. 2017*
- Integrated parameterized SQL views in Cognos Query Studio to allow clients to query user specific subsets of the data
- Produced multiple ad-hoc reports using SQL queries for providing detailed server and database metrics to management
- Wrote a PowerShell script to send customized emails according to configurable RDBMS stored information

Skills

Python	Good working experience with Python and familiarity with many libraries (Scikit-learn, Pandas/NumPy, Matplotlib, NLTK)
R	Intermediate level using R and familiarity using popular libraries (Tidyverse, Shiny, FlexDashboard, RStan)
Machine Learning	Good understanding and experience with classification, clustering, time series analysis and recommender systems
SQL	Experience with several SQL dialects (MS SQL, SQL Server and Oracle) and exposure to MongoDB and ETL tools such as Talend
Data Visualization	Good working knowledge of producing graphs using ggplot2 in R and matplotlib in Python; experience using Tableau
Statistics	Good applied statistics skills, such as hypothesis testing, regression analysis, time series analysis and Bayesian inference using STAN
Version Control	Familiarity using Git and some experience using Subversion
Languages	French (Fluent), English (Fluent)

Education

University of Ottawa *Ottawa, Ontario*
MASTER IN ELECTRONIC BUSINESS TECHNOLOGIES *Jan. 2018 - Exp. Dec. 2019*
- Research project: Text Mining for Intelligent Market Trading
- Relevant Courses: Data Mining and Concept Learning, Data Science Applications, Data Mining for Business, BI Technologies and Big Data

University of Ottawa *Ottawa, Ontario*
B.COM IN MANAGEMENT INFORMATION SYSTEMS AND ANALYTICS (MISA) MAGNA CUM LAUDE *Jan. 2014 - Dec. 2017*
- Received Merit Scholarship for maintaining an above average of 85 percent
- Relevant Courses: Statistics for Management, Applications of Statistical Methods in Business, Predictive Analytics

Projects and Certifications

Cannabis Recommender System   *July. 2019*

- Collected social media information (scraping Leafly reviews)
- Pre-processed and transformed the data (reviews using TF-IDF and one hot encoded categories) to build a content-based recommendation system
- Used the cosine similarity to provide recommendations and built an R Flex Dashboard to display the results
- Discovered two new restaurants that I loved

Text Classification and Clustering   *June. 2019*

- Used Python's scikit-learn and keras (feed forward neural network) to classify news articles
- Implemented clustering algorithms (EM and K-Means) to cluster the documents
- Evaluated the outcomes by calculating Kappa and Silhouette
- This will be my starting point for the Intelligent Market Trading research project

IMDB Movie Rating Prediction   *Apr. 2018*

- Built a parsimonious step-wise regression model to predict the Internet Movie Database (IMDb) movie ratings which was part of the Coursera course
- Extended my project by integrating my model into a Shiny application

2019   **Bayesian Statistics: From Concept to Data Analysis**, Coursera - Certificate received
2018   **Linear Regression and Modelling**, Coursera - Certificate received