

# Ryan Ondocin

+1(315) 345 3651 | [ryanondocin@gmail.com](mailto:ryanondocin@gmail.com) | [Data Science Portfolio](#) | <https://www.linkedin.com/ryanondocin> | Syracuse, NY

---

## Education

Syracuse University, School of Information Studies

May 2021

M.S. *Applied Data Science* | GPA: 3.9

Relevant Coursework: Big Data Analytics, Machine Learning and Data Mining, Cloud Management, Data Warehouse, Data Analysis and Decision

Making, Database Management, Business Analytics, Natural Language Processing, Information Visualization, Text Mining, Design Analysis of Algorithms

The Pennsylvania State University, University Park State College, PA

May 2019

B.S. *General Science, Minor in Chemistry*

## Technical Skills & Certifications

[Tableau Certified Desktop Specialist](#)

**Programming Languages and Cloud Platforms:** Python | SQL | R | Amazon AWS | DataBricks | Google Cloud Platform

**Software Tools:** Jupyter | RStudio | Microsoft SQL Server

**Libraries/Frameworks:** scikit-learn | NLTK | Mllib (Sparkml) | Apache Spark (PySpark) | TensorFlow | pandas | Numpy

**Visualization Tools:** Tableau | plotly | seaborn | matplotlib | Bokeh

## Experience

**Data Entry Specialist | State Fair Exposition Center** | New York State Department of Health

February 2021 – Present

- Organized and extracted electronic patient data for New York State's largest COVID-19 Vaccination Center
- Worked 12 hours/day, 4 days/week with a full courseload during final semester of graduate school
- Received special administrative privileges to manage patient appointments and forecast daily left over dosages
- Improved workflow efficiency by eliminating future survey questions based on previous selections

**Satellite Co-founder | Head of Developmental Operations** | Blackstone Launchpad

Nov 2019 – Present

- Employed Agile methodologies to manage a team of 6 Software Engineers to coordinate full stack development of a social media platform using React, Heroku, Amazon Web Services, PostgreSQL, and Ruby on Rails
- Facilitated bi-weekly meetings with developers to assign tasks and developed over 200 pages of documentation
- Participated in 5 Sharktank competitions, which granted the company \$5000 of funding to incorporate
- Named [BlackStone Launchpad's](#) Innovator of the Week, due to leadership qualities and management skills

**Machine Learning Research Assistant | PlantVillage Organization | Millenium Science Complex**

Jan 2017 – May 2019

- Trained object detection models to [help farmers diagnose plant diseases](#), offline, on their phones
- Annotated over 20,000 plant disease images to improve model accuracy by over 10 points to 93% via data augmentation
- Spearheaded collaboration between PlantVillage and IT Consulting firms, to coordinate front-end development, translations and debugging for mobile application
- Analyzed Satellite images via heat map analysis to monitor vegetation levels in African farm fields in 3 Countries

## Projects

**Clustering of Covid-19 Twitter Data** | Python | Machine Learning | NLP | SpaCy | NLTK | sklearn

May 2020

- Web scraped Twitter's API using tweepy to obtain 40,000 coronavirus related tweets
- Performed feature engineering and tokenization on textual data using the Natural Language Tool Kit
- Vectorized tweets via Tf-IDF and conducted Dimensionality Reduction via Principle Component Analysis
- Applied K-means and t-SNE to visually represent the transmission of COVID-19 tweets over a given time frame. An interactive Bokeh plot was created to investigate the the results [\[Interactive Plot\]](#)

**Disease Prediction** | Python | Machine Learning | sklearn | MICE imputation | plotly

March 2020

- Carried out EDA on a medical data set consisting of 50,000 rows, using statistical analysis and visualization techniques
- Implemented machine learning algorithms such as Decision Trees, Support Vector Machines (SVM), Random Forest, Naïve Bayes Classifiers, K-Nearest Neighbors (KNN), Logistic Regression and Artificial Neural Networks
- Replaced noisy and missing values using MICE imputation to obtain a 9% increase in recall
- Made predictions with 74% accuracy by utilizing GridSearch pipelines to conduct extensive model optimization

## Leadership and Community Engagements

**Band Leadership and Music Production**

Jan 2009 - Present

- Produced original songs on LogicProX and released to all streaming platforms, obtaining 26k plays on [Spotify](#)
- Created [Tableau](#) Visualization dashboards to analyze streaming sales and compare myself, auditorily, to other artists
- Performed at [Penn State's THON](#) event in front of over 200,000 viewers