

# DILLON DIATLO

## DATA SCIENTIST

✉ dillondiatlo@gmail.com 🌐 datadillon.netlify.app/ ☎ 732-492-3443 📍 New York City, NY  
in www.linkedin.com/in/dillondiatlo/ 📱 dillondiatlo

## SKILLS

**PROGRAMMING LANGUAGES:** Python, SQL, R, Scala

**TECH:** Git, Google Suite, Tableau, VS Code, Command Line, Databricks, Streamlit, Spark, Asana, Google Analytics

**SKILLS:** Data Cleaning, Data Pipelines, Algorithms, Statistical Analysis, Linear Models, Tree-based Models, Support Vector Machines, Clustering, Time Series, ARIMA, Neural Nets (DNN, CNN, RNN), Data Visualization & Graph Building, Data Analysis, Principal Component Analysis PCA, Natural Language Processing NLP, Data Mining EDA, Data Management, Web Scraping, Data Models, Recommendation Systems, Logistic Optimization, Marketing, Large Language Models

**PY LIBRARIES:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, BeautifulSoup, Tensorflow / Keras, PySpark

## PROJECTS

### NYC TAXI DRIVER PREDICTION TOOL

Apr. 2024

- Downsized 212M rows of data for statistical analysis, predictive and time series modeling
- A/B tested and feature engineered data pipelines for 5 different machine learning models
- Designed GradientBoostRegressor with 87% accuracy in predicting cost-per-trip, per-NYC taxi zone, by the minute
- Deployed predictive model with integrated OpenWeather API for real-time use via public [Streamlit app](#)

### NLP & SENTIMENT ANALYSIS | Reddit

Mar. 2024

- Scraped 10,000+ unstructured data points from subreddits r/running and r/swimming
- Built a structured dataset, cleaned, and conducted EDA and sentiment analysis
- Developed data pipelines to pre-process, fit, and test data with CountVectorization, TfidfVectorizer, lemmatizer, GridSearchCV, MultinomialNB models, RandomForest models, and more
- Designed a 97% accurate Logistic Regression model that can classify a subreddit post

### NEURAL NET SONG PREDICTION | Spotify

Mar. 2024

- Cleaned and analyzed 250,000+ rows of Spotify data to develop applied machine learning models including: RandomForest, DNN, LogisticRegression, and ExtraTreesClassifier
- Engineered a dense neural net model with 79% accuracy in predicting popular spotify songs
- Collaborated with team members via Git and articulated complex technical concepts for a non-technical 20 minute stakeholder presentation

### CNN IMAGE CLASSIFICATION

Mar. 2024

- Designed and implemented a convolutional neural net (CNN) on 4,200 image files using BigQuery and TensorFlow
- Co-created a Streamlit site for the CNN that can accurately recognize and classify a hotdog

## EXPERIENCE

### **CODE & THEORY, SENIOR COPYWRITER, New York, NY**

June 2022 - Aug. 2023

- Analyzed data to optimize client-website information hierarchies and UX modules
- Partnered with Project Coordinators to develop KPIs driven by industry data audits
- Wrote against client goals, KPIs, and metrics for full-copy revamp of Morgan Stanley, Simplify, and more
- Measured Amazon Prime Day advertising data to develop multiple Prime Day '23 campaign directions

### **R/GA, VERBAL DESIGNER, New York, NY**

Mar. 2020 - May 2022

- Researched and extrapolated data from 50+ client competitors to build strategic industry positioning
- Interpreted human-behavior data to design 10+ brand identities aligned with audience segments (brands ranged from Cigna's Evernorth to Dwayne Johnson's The XFL)

### **ENGINE GROUP, COPYWRITER & EDITOR, New York, NY**

Feb. 2018 - Jan. 2020

- Increased digital KPIs by 4000% through data-driven marketing campaigns
- Manipulated Gen-Z data to redesign the Birthright trip and win a \$500,000 client
- Sparked 32,000+ site visits in 90 days with B2B and B2C web copy overhaul

## EDUCATION

### **General Assembly**

Jan. 2024 - Apr. 2024

Data Science Certificate of Completion

700+ hrs of fully-immersive, hands-on data science program with 24 labs, 6 projects, and 6 presentations

### **University of Pittsburgh**

Bachelors of Arts in English

Aug. 2008 - Dec. 2012

## CERTIFICATIONS

Code Foundations Skill Path Certification, *Codecademy*

Nov. 2023

Learn How to Code Course Certification, *Codecademy*

Oct. 2023

How to Research and Write Using Generative AI Tools Certification, *LinkedIn*

Sept. 2023