DILLON DIATLO

New York City, NY · 732-492-3443 · dillondiatlo@gmail.com · [LinkedIn](#) · [Portfolio](#)

Machine Learning Engineer

Data Scientist with 10+ years of experience in the Advertising Industry.

EDUCATION

Data Science Certificate of Completion, General Assembly, NYC, NY   Jan 2024-April 2024
- 700+ hours of a fully immersive data science program with 24 labs, 6 projects, and 6 presentations

Bachelor of Arts in English, University of Pittsburgh, Pittsburgh, PA,   Aug 2008-Dec 2012

SKILLS

Languages: Python | SQL | R | Scala

Tech: Git | Google Suite | Tableau | VS Code | Command Line | Databricks

Skills: Web Scraping | Data Cleaning | Pipelines | Linear Models | Tree-based Models | Support Vector Machines | Clustering | Time Series | Neural Nets (DNN, CNN, RNN) | Data Visualization | Data Analysis | Principal Component Analysis (PCA) | Natural Language Processiong (NLP) | Data Mining (EDA) | Data Models | Data Management

Py Libraries: Pandas | NumPy | Matplotlib | Seaborn | Scikit-learn | BeautifulSoup | Tensorflow/Keras | PySpark

PROJECTS

PREDICTION TOOL NYC TAXI DRIVERS **,** *Streamlit, API, Tableau, VS Code*   March 2024
- Downsized 212M rows of data for analysis, predictive, and timeseries modeling
- Trained a GradientBoostRegressor with 87% accuracy in predicting trip location and cost by the minute
- Deployed predictive model via public Streamlit app for real-time use

SUCCESSFUL SONG FEATURES, JupyterLab, Github   March 2024
- Cleaned and analyzed 250,000+ rows of Spotify data and built RandomForest, DNN, LogisticRegression, and ExtraTreesClassifier models
- Trained a dense neural net model with 79% accuracy in predicting popular spotify songs
- Collaborated via Git with 3 data scientists for a 20 minute stakeholder presentation

NEURAL NETWORKS, Google Colab, Streamlit   March 2024
- Built and trained a convolutional neural network on 4,200 image files in 6 hours
- Co-created a Streamlit site for the CNN that can accurately classify an uploaded picture as a hotdog or not

NLP & SENTIMENT ANALYSIS, Reddit PRAW API, JupyterLab,   March 2024
- Mined 10,000+ unstructured data points from Reddit's r/running and r/swimming
- Built a structured dataset, cleaned, and conducted EDA and sentiment analysis
- Created pipelines to pre-process, fit, and test data with CountVectorization, TfidVectorizer, GridSearchCV, lemmatizor, MultinomialNB models, RandomForest models, and more
- Designed a 97% accurate Logistic Regression model that can classify a subreddit post

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
- Interpreted Amazon Prime Day data to develop multiple Prime Day '23 campaign directions

R/GA, VERBAL DESIGNER, New York, NY   March 2020 - May 2022
- Researched and extrapolated data from 50+ client competitors to build strategic industry positionings
- Interpreted human-behaviour data to design 12+ brand voices and identities, from Cigna's billion-dollar health services company Evernorth to Dwayne Johnson's The XFL, that would align with target-audiences

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