Jack Maney

Email: jackmaney@gmail.com

Website: jackmaney.com

- Data Science and Data Engineering professional with a decade of experience in the telecommunications, ad tech, financial service, and healthcare sectors.
- Able to derive insights and extract business value from data sets with thousands to billions of rows.
- Building prototypes of data products using relational databases, Apache Spark, MPP systems, and Python
- Working closely with Data Engineers to productionalize product prototypes.
- Researching and implementing algorithms from white papers and academic literature.
- Creating and A/B testing recommender systems.
- Experienced problem-solver with very strong quantitative skills.

Experience

American Century Investments

Senior Data Scientist - October 2023--October 2024

- Created and productionalized an Engagement Score for financial advisors.
- Took over maintenance of an existing production recommender system.
- Restructured parts of a Data Science monolithic repo, consolidating utility functions used in multiple places into a common utility library.
- Ran several ad-hoc analyses, some of which were under very strict deadlines.
- **Technologies Used:** Python, Amazon Redshift, pandas, scikit-learn, boto3, AWS, AWS Secrets Manager, AWS Lambda, S3, Tableau, git, GitHub

nbkc bank

Senior Data Engineer, AVP - November 2022--September 2023

- Did a nearly complete rewrite of the back-end for tracking movement of customers among segments. Using an object-oriented design, the addition of new segments or changes to existing segments are now much easier to implement.
- Built a heuristics-based approach for determining whether an ACH credit is a direct deposit.
- Altered Power BI dashboard of internal expenses to correctly show sub-account transactions. Assisted in showing results to the CFO and gathering feedback.
- Fixed incorrect assignments of Personal/Business accounts for survey data from customers.
- Automated a report used by the Deposits team.

- Used the Salesforce API to send emails when certain events are triggered.
- Mentoring members of the Data Team.
- **Technologies Used:** Python, pandas, Azure DevOps, Apache Airflow, Cosmos DB, Azure Key Vault, Power BI, git

Shopify

Data Scientist - June 2022--September 2022

Before being laid off, I was part of the Plans & Pricing Data Science team.

- Helped identify suspicious leads by implementing a markov chain to detect gibberish subdomain names.
- Took over the (manually run) referral program and worked with data engineers to automate this process.
- **Technologies Used:** Python, GCS, git

agilon health

Lead Data Scientist - July 2021--June 2022

- Worked on models to predict avoidable inpatient visits and medication adherence.
- Worked with the ML Ops team to validate datasets that have been moved to the Cloud.
- **Technologies Used:** Python, pandas, scikit-learn, S3, AWS SageMaker, git

Cerner Corporation

Data Scientist - July 2020--June 2021

- Used AWS SageMaker, BlazingText, and association rule mining to analyze and enhance the Chart Assist Ontology. Created a prototype that could make terminologists up to five times more efficient at matching medical codes to related ontology concepts.
- **Technologies Used: Python, pandas, XGBoost, AWS Sagemaker, BlazingText, git**

TruFactor

Lead Data Scientist - November 2018--May 2020

After Pinsight Media was acquired by InMobi, there was a reorganization into a few business units, one of which was TruFactor.

- Consolidated Point of Interest (POI) data from multiple sources to build a single POI database used across the company.
- Combined the POI dataset with GPS and network location data to build a Visits dataset. Worked closely with product managers to ensure that

the product aligned with customer expectations. Performed several white glove analyses for customer trials, which received positive feedback.
- Worked closely with Data Engineering to productionalize the Visits product, and to quickly deploy hotfixes and feature enhancements.
- Used multiple data sources--including scraping and third-party APIs--to build and maintain a mapping from URLs to publishers.
- Mentored members of the Data Science and Business Intelligence teams.
- **Technologies Used:** Python, PySpark, AWS EMR, AWS Lambda, S3, Snowflake, Tableau, H3, plotly, GeoJSON, git, GitHub

Pinsight Media

Data Scientist III - October 2016--October 2018

Pinsight Media was acquired by InMobi in 2018.

- Built recommender engines for our on-device monetization products. Worked with development teams to ensure that the back end would be set up for AB testing.
- Streamlined the ETL processes for building AB testing dashboards.
- Contributed to a project of predicting demographic attributes for users in our real-time bidding (RTB) system. This resulted in collaborations with Marketing that led to two whitepapers:
    - Guess My Age: Age Discrepancies Across Ad Exchanges
    - There's an App for That: But What Does It Say About My Consumer?
- Mentoring members of the Data Science and Business Intelligence teams.
- **Technologies Used:** Python, PySpark, AWS EMR, AWS Lambda, S3, Snowflake, Tableau, git, GitHub

DST's Applied Analytics Group

Senior Data Scientist - December 2013--October 2016

- Prototyped a product for acquisition of financial advisors.
- Contributed towards an Advisor Segmentation product, including a method of streamlining and summarizing the differences between segments.
- Built a prototype of the Mapper Algorithm (as used in Topological Data Analysis) to better understand high-dimensional data sets. The prototype is written in Python and leverages a Greenplum cluster by way of SQL templates.
- Built prototypes for three components of DST's Predictive Wholesaling product, and assisted the AAG Development team in productionizing the prototypes.
- Created and prototyped a Share Retention metric that provides a measurement of "stickiness" of fund holdings that does not directly depend on price.
- Assisted in building models for a proof of concept for a client.

- Mentored and taught Python to a few members of the Networking team, to facilitate the creation of a Flask web app to automate some types of network change requests.
- Mentoring other members of the Data Science team.
- **Technologies Used:** Python, Pivotal Greenplum, pandas, scikit-learn, git

BA Services

Data Scientist - May 2013--November 2013

- Created and cross-validated probit regression models to find the most significant attributes upon which to sort call queues in order to increase customer retention.
- Built an ETL pipeline to import data from a new dialer system.
- Delivered a proposal outlining options for a Data Warehouse solution, including pros and cons of each option.
- Built, Validated, and Deployed business intelligence reports using QlikView.
- **Technologies Used:** Perl, QlikView, SAS, SVN

C2FO

Implementation Specialist (Contract) - January 2013--May 2013

- Optimized the C2FO algorithm for Market Clearing Events, making it run an average of two orders of magnitude faster.
- Organized the restructuring of several KPI business intelligence reports.
- Built, tested, and deployed user management tools for account managers.
- **Technologies Used:** Perl, Ruby

Adknowledge

July 2010--November 2012
Titles Held:

- Sr Data Analyst and Mathematician - January 2012--November 2012
- Data Analyst and Mathematician - February 2011--January 2012
- Data Analyst - July 2010-February 2011

- Performed data mining and summarized results that contributed to the winning of a $50,000 advertiser contract.
- Presented technical and mathematical concepts to non-technical audiences, including several layers of management and a venture capital investor.
- Developed an application in Perl using DBI for k-means++ clustering. This application is able to handle data sets of millions of rows with 1--100 variables.

- Found a way to implement a regression algorithm--on a dataset with 30 million rows and 250 variables--that was previously thought impossible to implement due to scale.
- Built, implemented, deployed, and maintained an ad category recommendation system for advertisers, including developing and measuring performance metrics.
- Implemented a genetic algorithm framework to use for behavioral targeting algorithms.
- Maintained and documented the ETL pipelines to the Data Analytics team, consisting of over 200 scripts in Perl and Python interfacing with Greenplum, PostgreSQL, Oracle, MySQL, MS SQL, and ActiveMQ.
- Refactored and maintained critical business intelligence reports used by machine learning scientists.
- Prototyped a flexible, extensible ETL system to reduce a lot of boilerplate code in existing ETL scripts.
- Created a web-based data dictionary to store metadata about tables in our warehouse. The front end was written in PHP with SQLite on the back-end to store the metadata.
- Was considered a resident expert of our data warehouse.
- Contributed to the on-boarding of two interns and two full-time employees.
- **Technologies Used:** Perl, Pivotal Greenplum, PostgreSQL, Oracle, ActiveMQ, Python, PHP

University of South Dakota

Assistant Professor - August 2004--May 2008

- Six peer-reviewed mathematical publications.
- Directed two undergraduate Honors Theses and a Master's Thesis in mathematics.
- Sole organizer and director of a regional undergraduate mathematics conference.
- Taught several courses, including College Algebra, Trigonometry, Calculus (I--III), Foundations of Mathematics, Matrix Theory, and Abstract Algebra.
- Served and chaired several committees, including the Curriculum & Instruction committee.

Skills

Languages and Technologies

- Programming Languages: Python, SQL, Bash, R, Java, Kotlin, Scala, Perl
- Cloud Technologies:
    - AWS: Redshift, EMR, EC2, SageMaker, Lambda, Secrets Manager
    - Azure: Azure DevOps, Cosmos, Key Vault
    - Snowflake
- Apache Spark, Apache Hive, Apache Hadoop, HDFS

- Business Intelligence: Tableau, Microsoft Power BI, QlikView
- Python libraries of note: PySpark, BlazingText, H3, Shapely, Pandas, NumPy, SciPy, scikit-learn, Requests, matplotlib, seaborn, plotly
- Apache Airflow
- IDEs/Text Editors Used: PyCharm, IntelliJ IDEA, Visual Studio Code, Eclipse
- Data Formats: CSV, XSLX, JSON, GeoJSON, Parquet, XML
- Docker
- Version Control: Git, GitHub, GitLab, SVN, Assembla
- Task Tracking and Documentation: JIRA, Azure Devops Boards, Pivotal Tracker, Confluence, Assembla
- Operating Systems: Linux (Amazon Linux, openSUSE, Ubuntu, CentOS, RHEL), macOS, Windows, WSL, WSL2
- APIs used:
    - Salesforce (querying for specific events)
    - Zendesk (ETL on customer ticket data, adding custom fields)
    - SurveyMonkey (customer survey data)
    - Ramp (tracking internal expenses, creating card numbers on the fly)
    - Aylien (text analysis, NLP)
    - 42matters (mobile app metadata)
    - Infegy's Social Radar API (social media analytics)
    - Facebook Ads API

Other Skills

- Mathematics
- Python
- Topological Data Analysis
- Data mining
- Data visualization
- Implementing algorithms and ideas gleaned from academic publications

Education

North Dakota State University

Ph.D. Mathematics, May 2004

B.S. Mathematics, Dec 1999

Training Courses and Professional Development

- AWS MLOps Training and Workshop, 2024
- AWS Training from Amazon Web Services, 2018
- Apache Cassandra training from Learning Tree International, 2017
- Apache Spark training from Databricks, 2017
- Hadoop and MapReduce Training from Hortonworks, 2015

- Data Anonymization Training from Privacy Analytics, 2015
- Greenplum User Training from Pivotal, 2014
- Attended KDD 2014
- QlikView Developer Training from Qlik, 2013
- Noble Dialer Operations Training from Noble Systems, 2013
- Java Training from Webucator, 2012
- PostgreSQL Training from Webucator, 2010