

# Alexander Ilyin

650-504-8065 • [ailyin@ucsd.edu](mailto:ailyin@ucsd.edu) • <http://linkedin.com/in/alexander-ilyin1> • <https://alexilyin.me>

## EDUCATION

---

### UC San Diego | Masters of Science in Analytics (3.8/4.0)

June 2020

- Relevant Courses: Recommender Systems, Statistics for Analytics, Data Science for Marketing, Customer Analytics
- Recipient of Rady/UC Academic Fellowship, Admissions Ambassador

### UC Santa Cruz | Bachelors of Science in Management Information Systems – Honor's (3.5/4.0)

June 2019

- Relevant Courses: Calculus I-III, Linear Algebra, Probability Theory, Data Structures, Database Management Systems
- Dean's List, Member of Information Systems Management Association and Tennis Club

## SKILLS

---

**Programming Languages:** Python (Pandas, Matplotlib, Scikitlearn, Keras), R (ggplot, Plotly, Tidyverse), PostgreSQL, MySQL

**Data Visualization:** Tableau, R-Shiny, Python-Dash

**Other:** Python Web Frameworks – Flask/Django, PySpark/SparkR, AWS Sagemaker, ETL, A/B Testing

## PROJECTS

---

### MS Analytics Capstone Project

- Capstone Project team lead, organized project that involved using past social media interactions to optimize future social media campaigns for clients
- Used unsupervised clustering techniques to cluster brands based on their attributes, and evaluated social media success within clusters using a self-defined social media brand score
- Applied machine learning models in Python such as Logistic Regression, Random Forest, and Recurrent Neural Networks (RNN) to predict success of social media campaigns and products based on current trends
- Implemented natural language processing techniques such as topic modeling and sentiment analysis with visualization techniques such as word frequency and word clouds to allow exploration of social media trend fluctuations

### Twitter API Interface/Extension

- Built a Django-based interface for the Python Twitter API “Tweepy” to allow users to seamlessly scrape live data from Twitter
- Used the existing streaming capabilities of Tweepy paired with Celery task scheduling to run an asynchronous stream of tweets
- Tweets load directly to a backend PostgreSQL database, and the user gains access to a dashboard containing interactive visualizations and the option to extract tweets in CSV format for future analytics projects
- Users can also create Twitter chatbots using their own streamed tweets, or pre-trained chatbots (chatbots are created using RNN's in Keras)

## EXPERIENCE

---

### Data Analyst, Connor Group, Santa Clara, CA

July 2020 - Present

- Worked with clients to develop automated data pipeline solutions for internal financial systems to improve efficiency of legacy systems
- Met directly with clients' cross-functional finance, accounting, and engineering teams to gather and translate initial requirements, create documentation for planned and completed projects, and to present proposed pipeline solutions
- Performed thorough analysis and validation of data before creating requirements and data logic that was communicated back to clients to ensure that the proposed solution would meet both engineering and financial standards
- Developed financial data pipelines by API-based iPaaS solutions such as Workato and Dell Boomi. Pipelines followed the ETL process by extracting client's transactional data from different payment provider sources, transforming data to a standardized format, and loading data into either a data warehouse or an ERP software such as Netsuite
- Assisted the client with monitoring and debugging of the new pipelines following deployment

**Solutions Architect Intern,** Intrexon Corporation, South San Francisco, CA

June 2018 – August 2018

- Served as interface between engineering and experimentation teams to drive development of internal IT solutions
- Prototyped and deployed production database application to store trial and experimental data, optimizing the previous spreadsheet-based data storage method and creating a centralized data solution
- Led project through various aspects of the Agile product development lifecycle, including translating business requirements, creation of initial Entity Relationship Diagrams (ERD), and creation of a Minimum Viable Product (MVP) using a cloud-based database solution called Airtable
- New solutions allowed for automated data entry of experimental data as well as tracking of laboratory equipment
- Prototype application was used by team of scientists for storage of experiment results and inventory tracking