Justin Chen

chen.justi@northeastern.edu • 650-388-0985 • jyhchen.github.io/

EDUCATION

Northeastern University, Khoury College of Computer Sciences September 2019 - Present
Candidate for Bachelor of Science in Data Science and Business Administration (Fintech Concentration) Expected: Dec 2023
Awards and Honors: Dean's Scholarship, Dean's List, University Honors Program GPA: 3.54/4.0

Coursework In: Statistics; Object Oriented Design; no/SQL Information Storage; Supervised and Unsupervised Machine Learning; Data Visualization; Quantitative Financial and Economic Analytics
Involvement: Varsity Baseball Manager, Sports Analytics Club Advisor, WRBB Sports (student radio), Teaching Assistant

SKILLS AND INTERESTS

Computer Languages: Java, Python, R, Racket, HTML, CSS, Javascript
Database: SQL/mySQL, Google Big Query, Snowflake, MongoDB, Redis, Neo4j
Frameworks and Packages: Django, React.js, D3.js, AWS
Software: Git, Intellij, Juptyer, Adobe Analytics, Visual Studio Code, Tableau
Interests: Movie Soundtracks, Documentaries, Hats, Slow Pitch Softball, Biking
Other Skills: Social Media, Sports Writing and Broadcasting, Chinese (native speaker)

WORK EXPERIENCE

Data Analyst and Student Manager Boston, MA
Northeastern University Baseball January 2020- Present
- Managed team of student staff; oversaw and planned analytics and data initiatives
- Developed pitch and arsenal grades for game calling and player development with YakkerTech, and Trackman metrics
- Collaborated offense strategy with coaching staff to optimize value from lineups, bunts, and stolen bases
- Compiled cohesive scouting reports analyzing statistics, play-by-play, and video from Synergy and TruMedia

Quantitative Analytics Associate – Foundational Research (remote) Boston, MA
Philadelphia Phillies June 2023- September 2023
- Researched and analyzed the effects of airflow, spin vectors, and seams on a baseball's trajectory
- Utilized seam orientation data to quantify, model, and predict seam-shifted wake (SSW) movement
- Calculated spin axis pole locations, classified seam orientations, and identified unique pitch grips and releases
- Generated plots that map spin axes to assess a pitcher's arsenal and find optimal grips for SSW effects

Web Analytics and A|B Testing Co-op Brighton, MA
Harvard Business Review July 2022- December 2022
- Created interactive reports with Adobe Analytics to summarize A|B test findings and analyze user engagement
- Designed experiments to analyze article recommendations, newsletter engagement, and SEO article performance
- Developed model to predict website visitors' conversion to paying subscriber with 80% accuracy
- Automated dashboard backup and creation with R and Python, utilizing Adobe APIs

Baseball Systems Development Co-op Baltimore, MD
Baltimore Orioles July 2021- January 2022
- Rebuilt data pipelines for player stat views on the team's internal site with Django
- Developed amateur player evaluation forms and tables for scouts with React
- Designed quality checks and ID mapping for third-party data and created Slack notification bots for bad imports

PROJECTS

Stuff, Execution, and Prevention (Expected Run Values) Pitch Grader (Python/Jupyter Notebook)
- Modeled expected run values to quantify pitch metrics, like movement, velocity, release, in an aggregate
- Created pipelines for multi-parameter cross-validation and standardization to streamline preprocessing and training
- Evaluated feature importance and impact of different metrics to create the most comprehensive yet simplest models
- Graded both pro and amateur data to discuss the quality of pitches online and for Northeastern internal use

AirBnb Recommendation Engine (Python/Neo4j/MongoDB)
- Used graph database in Neo4j to create recommendations based on quantitative features and node distances
- Implemented engine in to query listings in MongoDB that match user inputs, then expand the search for similar listings
- Analyzed and discussed the advantages and shortcomings of each database at accomplishing similar tasks