

# Yao Yao

U.S. Citizen, Available for Relocation, Seeking challenging position in data science

Project Summaries: <https://www.linkedin.com/in/yaoya0/>

Projects: <https://github.com/yaowser/>

1417 W Fillmore St. Apt 3E

Chicago, IL 60607

(248) 602-0798

yao.y89@gmail.com

---

## EDUCATION & AWARDS

Southern Methodist University, **Masters of Science in Data Science**, Focus in ML, GPA: 4.0 Aug 2018

Oakland University: Completed MBA courses in business analytics, finance, statistics May 2016

University of Michigan, **Bachelors of Science in Material Science Engineering**, GPA: 3.36 May 2012

**Awards:** Second Place for Capstone Poster (SMU), Awarded \$25,000: Michigan Clean Energy Challenge (Warmilu)

## TECHNICAL AND LANGUAGE SKILLS

**Software:** Python, R, Databricks/Spark, Scala, Tableau, SQL, C++, SAS, VBA, Java, Keras, Tensorflow

**Languages:** English (first language), Mandarin Chinese (fluent), German (working proficiency)

**Certifications:** R, Python, SQL, machine learning, predictive analysis, linear algebra, Processing 3, SQL, C++, HTML, CSS, Probability, Statistics, VBA, and SAS via Udemy, Udacity, EDX, Khan Academy, and Codecademy

## EXPERIENCE

### **Graduate Student Researcher: 11 Major Projects + Capstone (Published)**

Dallas, TX

*Southern Methodist University*

Jan 2017 – Aug 2018

- Team leader, writer, and presenter for dataset, problem, workflow methods, interpretations, and applications
- Yelp's Review Filtering Algorithm (Thesis), ML Audio Source Separation, Taught Databricks & Blockchain
- Mastered: optimization, classification, feature importance, validation, regression, visualization, imputation, proportional sampling, reduction, LDA, PCA, ARIMA, clustering, Naïve Bayes, Stanford NLP, decision trees, random forest, XGBoost, SVM, neural networks, Hidden Markov, OOP, polymorphism, data scraping

### **Graduate NLP and Semantics Researcher**

Chicago, IL

*Univ. of Illinois Chicago Big Data Symposium/CRIM*

Mar 2016 – Dec 2017

- Researched new product and market creation by analyzing large scale, online interactions on Reddit
- Parsed unstructured data into relational databases by applying NLP and semantic tagging via Python

### **Product Technology – Management Associate**

Troy, MI

*U. S. Steel*

Sept 2012 – Dec 2014

- Automated spreadsheets and summary reports for lab data, revenue, and shipping logistics using VBA
- Created Java forms to enhance data entry at manufacturing plants to track scraps, rejects, and costs
- Developed new geometries for existing parts, accompanied by progressive stamping methods using CAD
- Investigated the formability, weight distribution, and costs of steel geometries with that of aluminum

### **Co-founder, Design and Test Engineer (Simulations)**

Ann Arbor, MI

*[Warmilu](#), Warming Blanket for Premature Infants*

Sept 2011 – May 2012

- Evaluated material heat transfer, heat insulation, and price optimization for blanket prototype fabrication
- Simulated thermodynamic design using 3D FEA to keep hypothermic body temperature at 37°C for 4 hours
- Ran thermocouple testing with LabView to verify heat transfer and retention rates of working prototype
- Developed a blanket that complies with medical regulations in accordance to Mott's Children's Hospital

### **Undergraduate Solar Cell Research**

Ann Arbor, MI

*Polymer Solar Cell Synthesis with Professor Jinsang Kim, MSE Department*

Sept 2010 – May 2012

### **Undergraduate eBay Research ([Abstract Published](#))**

Ann Arbor, MI

*Behavioral Game Theory on eBay with Professor Romesh Saigal, IOE department*

Sept 2009 – Aug 2010

- Wrote programs in VBA and MATLAB that imported user data from eBay's API into support vectors
- Derived perceived values of all eBay items using Nash Equilibrium and Game Theory decision trees
- Formulated the seller reliability index of every eBay user based on seller attributes with machine learning