# QUALIFICATIONS SUMMARY

I am highly analytical and versatile professional with over nine years of industry experience in development of predictive modeling applications with large datasets. I have strong background in computer sciences, statistics, and machine learning approaches and in application of these subject areas in research and commercial setting.

* Experienced in developing commercially successful prediction products and in managing large data research projects involving qualitative, quantitative, and complex response variables and attribute sets.
* Leveraged big data tools in research setting to provide solutions to process large amounts of genomic data and perform analytics.
* Experience in programming with Python, R, SAS, SQL, C++, Unix Shell, and PHP. Experience working in Grid and AWS environments.
* Actively administered and participated in software development for commercial applications within organization.
* Bachelors in Computer Science, Masters in Computer Science, PhD in quantitative and statistical Genetics.

**PROFESSIONAL EXPERIENCE**

# Principal Data Analyst

## Capital One Home Equity, Plano, TX Aug 2016–Present

## 

* Lead the data analyst team for risk, compliance and SLQ; build data ecosystem of real-time origination process control.
* Work with software engineer for building the new generation of loan fulfillment UI as SME and business logic provider; work with data engineering to build Cassandra API for workflow management; solve any data related problems and monitor data accuracy.
* Major contributor to Home Loan Spamhuas/Whitelist project; pioneered, designed and automated data process for origination email collection, comprehensive validation and data quality control, which was set an example and standard for other LOBs by the enterprise team.

Award: Capital One Circle of Excellence, Q3 2016

# Senior Data Analyst

## Capital One Home Loan Operation, Plano, TX Feb 2014-July 2016

* Supported a full spectrum of operation data needs: underwriting, loan processing, quality assurance, customer experience, compliance and audit.
* Designed and developed data products and efficiency tools (Python and Tableau) for 200+ loan processors, underwriters and team leaders; helped to improve underwriting SLA adherence from 73% to 90% in three months; helped home equity to double funded loans per FTE in 6 months.
* Managed and coordinated the complete life cycle of data products – from project initiation, empathic design, porotype development, floor test and feedback collection, result monitoring and agile improvement.

Awards: Financial Service Mark of Distinction, Q1 2016

First Place, Capital One 2015 Data Community All Hands Expo Competition, Oct 2015

# Data Analyst Consultant / Senior Data Analyst

## Capital One Home Loan Marketing & Analysis, Plano, TX Jan 2012–Feb 2014

* Lead data analyst to support Home Loan Marketing & Analysis business and statistician teams for loan origination, pricing, credit policy and valuation
* Worked with statisticians to design and build the data wrangling process to wrap credit, origination, and loan performance data from different systems and sources into a consolidated database for new application model building.
* Debugged, streamlined and redesigned the model monitoring and quality control process for home equity account level valuation (ALV) models.
* Created standardized scripts to build data warehouse for analysis purpose using SAS/SQL in UNIX environment with Teradata as database; Build ad hoc reports for loan fulfillment process monitoring using SQL and Business Objects.

Awards: Home Loan Circle of Excellence, Q1 2014

Financial Service Statistician Team STAR Award

# Research Assistant/Associate

## Iowa State University, Ames, IA Aug 2006- Dec 2011

* Applied statistical learning and data mining methodologies to engineering materials design
* Modeled the nonlinear composition-processing-property relationships of alloys using recursive partitioning, principal component analysis (PCA), artificial neural network(ANN), and genetic algorithm (GA)
* Performed 3-way ANOVA statistical analysis on nanoindentation hardness data of metal thin film
* Used statistical software and languages: Matlab, JMP, SAS, R

## Publication and Patent: <https://scholar.google.com/citations?user=mZ0UPNEAAAAJ&hl=en>

# EDUCATION

Master in Statistics: Iowa State University, expected summer 2017

Ph.D. in Engineering: Iowa State University, Ames, IA

Bachelor in Engineering and Computer Science (dual degree): Xi’an Jiaotong University (XJTU), Xi’an, China

## 

# RELEVENT COURSES

*Iowa State University*

STAT500 Statistical Methods, STAT501 Multivariate Statistical Methods, STAT510 Statistical Methods II – Linear Models, STAT520 Statistical Methods III, STAT531 Quality Control and Engineering Statistics, STAT542 Probability Theory, STAT543 Statistical Inference, STAT579 Statistical Computing with R, STAT511 Time Series Analysis, IE583 Knowledge discovery and data mining, CS573 Machine Learning (audit)

*Xi’an Jiaotong University, China*

Fundamental of Database Systems, Data Structures and Algorithms, Fundamental of Operating Architecture, Principles of Computer Network