**Yaotian Zou**

zouyaotian@hotmail.com

979-739-3681

100 Station Dr., Avenel, New Jersey, USA, 07001

**SUMMARY**

* 3 years of professional data science / predictive modeling / analytics experience. Worked closely with product managers, actuaries, and marketing specialists to propose and apply data-driven solutions in auto insurance industry (risk-based premium ricing, marketing, underwriting, customer retention, etc.).
* 7+ years of industrial and academic practice in advanced statistical modeling and machine learning techniques such as linear regression, generalized linear models (GLM), classification and regression trees (CART), random forests, generalized boosted models (GBM), support vector machine (SVM), ensemble modeling, Bayesian methods and unsupervised clustering methods.
* 8+ years of experience in advanced use of R, SAS, Python and SQL for machine learning, feature engineering, visualization and data manipulation.
* 5+ years of outstanding academic research (master’s, doctoral, and postdoctoral) in real-world application of advanced quantitative methods with multiple publications and conference presentations.
* Currently on H1B working visa.
* Willing to relocate.

**EXPERIENCE**

**Data Scientist,** 11/2015 - Present

Plymouth Rock Assurance, Red Bank, New Jersey, USA (medium-sized auto insurance company)

* Developed a ground-up vehicle ranking model for pricing the personal injury protection (PIP) insurance coverage based on vehicle characteristics and safety features. Improved the premium segmentation by 10% as a pioneer project in auto insurance industry.
* Built marketing response, quote, sold and yield models for optimizing leads purchasing based on third party demographics and credit data. Improved marketing campaign’s yield rate by 20% to 30%.
* Proposed a customer retention model which allows predicting customers’ life time value at different point of time. Substantially optimized marketing expense strategies in new business acquisition.
* Developed a proprietary model based on insureds’ prior carrier data (e.g. length of insurance, lapse in coverage, etc.) which helped both pricing enhancement and underwriting.
* Developed an unconstrained model for auto pricing and presented its performance metrics against current pricing model to executive management team on a quarterly basis. Automated the process from data preparation to results evaluation which reduced the work load by 75%.
* Provided technical support on model implementation, maintenance and updates.

**Research Assistant (Ph.D.)**, 08/2012 - 10/2015

Center for Road Safety, Purdue University, West Lafayette, Indiana, USA

* Performed advanced statistical and econometric analysis on motor vehicle crashes to aid decision making in evaluating and proposing road safety treatments for Department of Transportation (DOT) engineers.
* Assessed the change in injury risks and crash frequencies due to the use of road barriers using GLM regression methods and provided Indiana DOT with strategic alternatives and numerous scenarios for optimal use of different types of barriers, and the results were published in a peer-reviewed journal.
* Provided valuable insight by leveraging multiple analytical and statistical techniques including exploratory data analysis (EDA), linear and non-linear regressions, decision trees, clustering methods, statistical simulation, machine learning and predictive modelling methods.
* Managed the full cycle of analysis including writing proposal, collecting and processing data, programming, analytical and statistical modelling, presenting and tracking the results.

**Graduate Student (M.S.)**, 08/2010 - 07/2012

School of Civil Engineering, Texas A&M University, College Station, Texas, USA

* Developed a new count model based on the double Poisson distribution and examined its potential application in modelling the motor vehicle crash frequency.
* Programmed its parameter estimation algorithm based on both maximum likelihood estimation and Bayesian methods (Markov Chain Monte Carlo) using R software. The results were presented in a national conference and published in a peer-reviewed journal.

**EDUCATION**

**Ph.D.** in Civil Eng., Purdue University, West Lafayette, Indiana. (GPA: 3.97/4.0) 08/2012 - 12/2014

**M.S.** in Civil Eng., Texas A&M University, College Station, Texas.(GPA: 4.0/4.0) 08/2010 - 08/2012

**B.E.** in Transportation Eng., Southeast University, Nanjing, China. (GPA: 87/100) 08/2006 - 06/2010

**AWARDS**

2017 President’s Award of Plymouth Rock (top 5% employee based on annual performance review)

**PUBLICATIONS**

**Zou, Y.** andTarko, A. P. (2018). Barrier-relevant crash modification factors and average costs of crashes on arterial roads in Indiana. *Accident Analysis & Prevention*, Vol. 111, pp. 71-85.

**Zou, Y.** andTarko, A. P. (2016). An insight into the performance of road barriers - redistribution of barrier-relevant crashes. *Accident Analysis & Prevention*, Vol. 96, pp. 152-161.

**Zou, Y.,** Tarko, A. P., Chen, E., and Romero, M. A. (2014). Effectiveness of cable barriers, guardrails, and concrete barrier walls in reducing the risk of injury. *Accident Analysis & Prevention*, Vol. 72, pp. 55-65.

**Zou, Y.,** Geedipally, S.R. and Lord, D. (2013). Evaluating the Double Poisson Generalized Linear Model. *Accident Analysis & Prevention*, Vol. 59, pp. 497-505.

**CONFERENCE PROCEEDINGS**

**Zou, Y.,** Geedipally, S.R. and Lord, D. (2013). Evaluating the Double Poisson Generalized Linear Model. Presented at the 92nd Annual Meeting of the Transportation Research Board, Washington. D.C..

**Zou, Y.**, Lord, D. and Geedipally S.R. (2012). Over- and Under-Dispersed Count Data: Comparing the Conway-Maxwell-Poisson and Double-Poisson Distributions. Presented at the 91st Annual Meeting of the Transportation Research Board, Washington. D.C..