

Roger Yeh

Manager, Data Science - Experian Consumer Services

Irvine, CA - Email me on Indeed: [indeed.com/r/Roger-Yeh/47fd14477e0a7097](https://www.indeed.com/r/Roger-Yeh/47fd14477e0a7097)

A result oriented and data driven analytical professional focusing on turning data into actionable insight.

RELEVANT MAJOR PROJECTS

- Marketing Mix Model: Provide a holistic view of mix channel marketing measurement through integrating Attribution Portfolio, Marketing Efficiency, and Acquisition Order Forecast
 - CLV based consumer Segmentation: Hybrid Survival Tree approach to target consumer profile based on their lifetime value
 - Analytics Framework: Architect and implement comprehensive predictive insight cycle, including model deployment for real time scoring in operational state
 - Recommendation Engine: Develop logistic regression modeling to generate content based article recommendation based on Semantic Parsing/Collaborative Filtering distance scoring
 - IoT: Utilized signal-processed sensor data combined with principal component reduction to predict wind turbine failure rate
 - Multiple Regression Analysis: Created multiple general linear regression model to forecast used/new/leased vehicle residual
- Authorized to work in the US for any employer

WORK EXPERIENCE

Manager, Data Science

Experian Consumer Services - February 2016 to Present

- Marketing Mix Model
 - Attribution Portfolio
 - Channel Efficiency
 - Acquisition Forecast
- Customer Lifetime Value Based Segmentation
 - Apply survival tree to segment consumer profile based on lifetime value
 - Provide a decision rules to target on specific consumer profile and resulted significant increase on customer retention and ROI
- Analytical Framework Architecture Design
 - Design an operational predictive model deployment platform on Google Cloud and Amazon Web Services
 - Real time actionable insight helps the business to increase ROI and monitor daily business performance

Senior Data Scientist

Predixion Software - February 2015 to February 2016

- IoT(Internet of Things)
 - Predictive Maintenance: Predict failure rate on oil well/windmill turbine
 - Asset Management: Tire failure risk on fleet of trucks
- Healthcare
 - Biologics treatment classification based on Psoriasis patient(Novartis)
 - Work Center Turn Around Time Classification and Experimental Design for campaign effectiveness(Athena Health)

Senior Quantitative Analyst

J.D. Power - September 2014 to January 2015

- Vehicle Valuation For Total Loss Vehicle
- Disaster Recovery Server Migration on AWS and UAT
- Insurance Practice: Various Ad-Hoc to support client request on total loss vehicle

Data Scientist

- April 2012 to June 2014
- Recommendation Engine Modeling and Amazon Big Data Implementation
 - Increasing web traffic by 4% visit and 3% ad click rate
 - Increasing magazine subscription rate and attract new subscriber from referral recommended
 - SAS and R deployment on AWS instance resulted improved modeling score process by 150%
- Cross Channel Subscriber Behavior Analysis
 - Data feed through Pixel Tracking Web Logs hosted in AWS S3
 - Provide strategic insight based on cross website behaviors
 - Subscription campaign driven by advanced subscriber profiling
- US Auto Sale Industrial Insight(Quarterly Auto Sale forecast for pay per click campaign)

Senior SAS Programmer

UnitedHealth Group - September 2011 to April 2012

- Program optimizing and integrate internal Risk-Adjust claims data structure layout and report generating efficiency, cutting down several production processing time by at least 50%
- Develop Projection Forecast on 2012 high risk-patient call campaign(reduce unnecessary call expense for need to care campaign)

Statistician

Kelley Blue Book - October 2007 to September 2011

- New/Used Vehicle Pricing Forecast - Using General Linear Regression to sustain accuracy at 6% error on new/used car price forecast
- Classification on Mileage segment groups based on two stage segmentation: Principle Component and Predictive Cluster Analysis
- Risk Assessment Model on Brand Demise(Oldsmobile/Saturn/Mercury)

- Production Data Optimization - by various SAS data manipulation technique:
- Decrease the processing time of data import procedure by 80% by deploying Netezza analytic data warehouse
- Full process automation to eliminate analyst interference and process failures

Consultant

Statistical Consulting Center - January 2007 to September 2007

- Conducted analysis of treatments for sleeping disorder by applying complete statistical analytical cycle
- Conducted analysis of treatments for swallowing disorder with dysphasia treatment
- Used Six Sigma method to improve CSULB bookstore efficiency and productivity

Scientifics Platform and Programming

- Data Science & Modeling: SAS 9.4, R Python
- Analytics: Google Analytics 360, Omniture - Site Catalyst/Report Builder/Data
- Big Data Infrastructure: Warehouse SQL/MySQL (Relational), Amazon EMR/Redshift(Unstructured)
- Reporting and Visualizations: Spotfire Reporting System, Tableau

EDUCATION

Master of Science in Applied Statistics

CALIFORNIA STATE UNIVERSITY

August 2007

Bachelor of Science in Mathematics

UNIVERSITY OF CALIFORNIA - Irvine, CA

September 2003