CESAR ACOSTA

cesar-acosta.github.io | linkedin.com/in/cesar-acosta | cesar.acostamejia@gmail.com

More than 10 years of experience analyzing highly complex data, building advanced machine learning models to predict market outcomes useful to improve decision making in Marketing Analytics, Financial investing, and business operations analytics. Outstanding skills to finding insights for problem solving and process optimization.

EDUCATION

- Ph.D. Statistics, The University of Texas at Dallas
- M.Sc. Statistics, The University of Texas at Dallas
- Ph.D. Industrial Engineering, Texas A&M University
- MBA, Monterrey Tech (ITESM) México City Campus

SKILLS

- Predictive Analytics: Machine Learning - Python
- Prescriptive Analytics with Simulation
- Data Mining with Python, R
- Deep Learning with Python
- Advanced Statistical Analysis
- Data Visualization: *R* ggmap, Tableau

Ample experience applying Machine Learning to solve real-world problems through the use of advanced statistical analysis to build prediction and classification models in applications such as Marketing Research, Portfolio Optimization, Volatility Forecasting, and Automated trading. Experience using Deep learning, Ensemble Methods, Gradient boosting, and other classification methods (KNN, Discriminant Analysis), Regression methods (Ridge, LASSO, Logistic, Multinomial regression), and Clustering methods (PCA, K-means, Hierarchical clustering).

PROJECTS

- A marketing strategy required to group a large set of US metropolitan cities into clusters. It was deemed critical to characterize the cities in each of the clusters. A K-Means algorithm was used to cluster the cities and a biplot and pivot tables were used to characterize the cities in each cluster. The project was completed with python. It was found which attributes the cities in each cluster have in common and how much the cities from different clusters are distinct. This solution helped the Director to efficiently allocate resources and marketing efforts to the different clusters attaining the best possible revenue per dollar invested. March 2024.

- A health insurance company asked for help to estimate the medical expenses charged to their top insurance plan for the next year. The manager wanted to predict how much more expensive is the premium for each additional children and how much the medical expenses increase for each additional year of age. To this end a nonlinear regression model is developed using the patient's age, sex, body mass index (bmi), number of children, smoker (yes or no), and location as predictors. Some useful interactions are found that improve the model accuracy. The results allowed the company answer many business questions such as how much more expensive is a smoker than a non-smoker, and how much more expensive are obese smokers than non-smoker patients that are not obese. August 2023.

- The client customer data from a large number of emails is used to predict if they are spam. Classification models were built to find the most accurate predictions. 57 variables indicating the frequency of certain words in the email were used in the models. It is expected that these words are useful for prediction. 5-fold cross validation was used to estimate and compare the models performances. The best model attained a 94% test accuracy rate. May 2022.

- Mixtures-based Value at Risk Estimates of Financial Stocks. Value at risk (VaR) and Conditional VaR (CVaR) are two common measures of risk that are related to the loss distribution. It is generally believed that if the true loss distribution is heavy-tailed, as compared to the normal, then the risk is higher. It was shown that in general this is not the case. Formulas for VaR and CVaR for mixtures were derived and it was shown that there are instances where the normality assumption overestimates (and the mixture distribution underestimates) the observed market risk. January 2020.

- Improved an S&P 500 Investment tracking portfolio by changing the assets allocation to reduce the associated risk. Investment portfolios with and with no assets' shortselling were derived. An optimization model with cardinality constraints was considered for each case. The portfolio's risk was reduced in term of the return's volatility. Portfolios based on other risk measures such as inverse volatility portfolio, equal-risk-contribution portfolio, and, maximum diversification portfolio, were considered. December 2019.

- Compared the performance of Random Forest with Multinomial Regression to classify customers into several market segments. It was shown that Random Forest and Ensemble methods outperform multinomial regression models in terms of predictive performance. July 2018.
- Optimized Marketing Resource allocation to improve customer acquisition and retention. The nonlinear relation between marketing efforts and customer acquisition and retention was derived. Values of annual spending on acquisition and retention that maximize return on investment were found. July 2017.

EXPERIENCE

- Data Scientist Consultant *2013 - 2024*
- USC, Viterbi School of Engineering, MS Analytics, program director *2018 - 2024*
- USC, Viterbi School of Engineering, MS Analytics, Predictive & Prescriptive Analytics, instructor *2016 - 2024*
- USC, Viterbi School of Engineering, MS Analytics, Data Mining, instructor *2016 - 2024*
- USC, Viterbi School of Engineering, MS Financial Engineering, instructor *2014 - 2018*
- USC, Department of Industrial & Systems Engineering, BS instructor *2014 - 2018*

INVITED SPEAKER

- IDEAS 2019 Conference on AI. *Data Science and Analytics. Competing in a data-driven World.* International Data Engineering and Science Association. October 2019, Los Angeles, CA.
- SatRday LA 2019 *Multiple Response Regression Models.* Los Angeles *R* Users Group. April 2019, Los Angeles, CA. `https://losangeles2019.satrdays.org/`
- IDEAS 2018 Conference on AI. *Is the Best Predictor actually the best?.* International Data Engineering and Science Association. October 2018, Los Angeles, CA. `www.ideassn.org/socal-2018/`

PUBLICATIONS

- Acosta-Mejia, C. A., Rincon, L. A., "The Continuous Run Sum chart", *Communications in Statistics - Theory and Methods*, 43: 4371 - 4383, 2014.
- Acosta-Mejia, C. A., "Two-sided charts for monitoring nonconforming parts per million", *Quality Engineering*, 25, pp. 34 - 45, 2012.
- Acosta-Mejia, C. A., "On the Performance of the Conditional Decision Procedure in Geometric charts", *Computers and Industrial Engineering*, 61, pp. 905 - 910, 2011.
- Acosta-Mejia, C. A., Pignatiello J. J., "The Run Sum R chart with fast initial response", *Communications in Statistics - Simulation and Computation*, 39, pp. 921 - 932, 2010.
- Acosta-Mejia, C. A., Pignatiello J. J., "ARL-Design of S Charts with k-of-k Runs Rules", *Communications in Statistics - Simulation and Computation*, 38, pp. 1625 - 1639, 2009.
- Acosta-Mejia, C. A., Pignatiello J. J., "Modified R charts for improved performance", *Quality Engineering*, 20, pp. 361 - 369, 2008.
- Acosta-Mejia, C. A., "Two sets of runs rules for the $\overline{X}$ chart", *Quality Engineering*, 19, pp. 129 -136, 2007.
- Acosta-Mejia, C. A. Pignatiello J. J., "Monitoring the Variability of Symmetric Processes", *International Journal of Industrial Engineering*, 9, pp. 151-161, 2002.
- Acosta-Mejia, C. A., Pignatiello J. J., "Monitoring Process Dispersion with no Sub-grouping", *Journal of Quality Technology*, 32, pp. 89-102, 2000
- Acosta-Mejia, C. A., Pignatiello, J. J., Rao, V. B., "A Comparison of Control Charting Procedures for Monitoring Process Dispersion", *IIE Transactions*, 31, pp. 569-579, 1999
- Acosta-Mejia, C. A., "Improved p charts to Monitor Process Quality", *IIE Transactions*, 31, pp. 509-516, 1999
- Acosta-Mejia, C. A., "Monitoring Reduction in Variability using the Range", *IIE Transactions*, 30, pp. 515-523, 1999

AUTHORED BOOK

Financial Derivatives, 2018. My textbook for a course in financial derivatives, portfolio optimization, and hedging. The book includes examples and exercises in R to construct optimal portfolios, to estimate Value at Risk, to price European and American options, among other applications. It also introduces Stochastic processes and stochastic calculus for the Black and Scholes formulas, and covers Monte Carlo simulation of Brownian motion to estimate the price of some exotic options. It shows how to use libraries `RQuantlib, Rmetrics, rugarch, fOptions, fExoticOptions` for financial modeling.

AWARDS

- 2018-2019 Outstanding Teacher of the Year, USC Department of Industrial and Systems Engineering
- 2018, Best FE Track Paper Award. Castro R., Huang S., Liu J., Blay R., Acosta-Mejía C. *Mixtures-based Value at Risk Estimates of Financial Stocks*. Third North American International Conference on Industrial Engineering and Operations Management, IEOM Society International.

TEACHING

Courses offered with average teaching ratings (in parentheses)

- ISE 529 Predictive Analytics, MS Analytics, 2023, USC (4.51/5.00)
- ISE 535 Data Mining, MS Analytics, 2024, USC (4.74/5.00)
- ISE 580 Prescriptive Analytics with Simulation, MS Analytics, 2023, USC (4.80/5.00)
- ISE 563 Financial Engineering, MS Financial Engineering, 2014 - 2018, USC (4.57/5.00)
- ISE 225 Engineering Statistics, BS ISE, 2015 - 2018, USC (4.38/5.00)
- ISE 220 Probability Concepts for Engineering, BS ISE, 2014 - 2019, USC (4.41/5.00)

UNDERGRADUATE/GRADUATE RESEARCH

- Castro R., Huang S., Liu J., Blay R., Acosta-Mejia C. *Mixtures-based Value at Risk Estimates of Financial Stocks*. Best Track Paper Award at the Third North American International Conference on Industrial Engineering and Operations Management, IEOM Society International, September 2018, Washington, D.C.
- Wang Q., *Multiple Response Regression Models using R*. Presented at the 2019 Saturday *R* day in Los Angeles. Wang is an Analytics MS student supervised by Cesar Acosta.
- Kim S., Upadhyay S., Acosta-Mejia C. *Using Categorical Variables in Predictive Analytics*, Annual IISE Conference and Expo, May 2018, Orlando, FL.
- Munoz N., Collins T., Acosta-Mejia C. *Identifying Outliers in Large Mixed Databases*.
  presented at the Annual IISE Conference and Expo, May 2016, Anaheim, CA.
  Munoz and Collins, ISE undergraduate senior students supervised by Professor Acosta.
- Lin Z., Chen Z., Acosta-Mejia C. *Robust Multivariate Control Charts*.
  presented at the Annual IISE Conference and Expo, May 2016, Anaheim, CA.
  Lin and Chen, ISE and OR masters students supervised by Professor Acosta.

PRESENTATIONS

- Kim S., Upadhyay S., Acosta-Mejia C., Using Categorical Variables in Predictive Analytics, Annual IISE Conference and Expo, May 2018, Orlando, FL.
- Castro R., Huang S., Liu J., Acosta-Mejia C., Mixtures-based Value at Risk Estimates of Financial Stocks, 3rd. North America IEOM Conference, September 2018, Washington, D.C.
- Munoz N., Collins T., Acosta Mejia C., Identifying Outliers in large mixed databases. Annual Industrial and Systems Engineering Conference and Expo, Institute of Industrial Engineers, May 2016, Anaheim, CA
- Lin Z., Chen Z., Acosta Mejia C., Robust Multivariate Control Charts. Annual Industrial and Systems Engineering Conference and Expo, Institute of Industrial Engineers, May 2016, Anaheim, CA
- Acosta-Mejia, C. A., Yarto G. J., The Modified Run Sum chart. Annual Industrial and Systems Engineering Conference and Expo, Institute of Industrial Engineers, May 2010, Cancun, Mexico
- Acosta-Mejia, C. A., The Run Sum R chart. Annual Industrial and Systems Engineering Conference and Expo, Institute of Industrial Engineers, May 2010, Cancun, Mexico

ENGINEERING INTERNS

- Juan Diaz, BS ISE Student. Intern at Microsoft Co., 29 May 2018 to Aug 17, 2018. *Project Manager Assistant.*
- Shaligram Tanvi, MS ISE Student. Intern at Energy Recovery Inc., Jan 22 to May 4, 2018. *Project Manager Assistant.*
- Yue He, MS ISE Student. Intern at WirelessBro Inc., Jan 8 to May 11, 2018. *Growth Marketing Analyst.*
- Varunya Ilanghovan, MS NLTX Student. Intern at Siemens PLM Software. June 11 to Aug 17, 2018. *Data Analyst.*
- Kanagarajan Shanjay, MS ISE Student. Intern at Faraday Future Inc., *Faraday Future 01 Beta Vehicles*, Daniel J. Epstein Department of Industrial and Systems Engineering, Spring 2017.
- Sinuo Song, MSFE Student. Intern at Amerihome Mortgage Co., *Customer segmentation*, Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2017.
- Upadhyay Sudhanshi, MS ISE Student. Intern at Applied Medical Co., *Product Lifecycle Management*, Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2017.
- Shaligram Tanvi, MS ISE Student. Intern at Energy Recovery Inc., *Quality Engineering - Ceramic Lab*, Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2017.
- Lakshman Supritha, MS EM Student. Intern at Mindtree Ltd., *Application Management services (AMS) with S/4 HANA*, Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2017.
- Liu Jiajian, MS ISE Student. Intern at D.F. Stauffer Biscuit Co., *Manufacturing process improvement project*, Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2017.
- Korada, J. Directed Individual/Independent study at Internet Brands Co. *Migration of iControl payment and billing system to Salesforce* Daniel J. Epstein Department of Industrial and Systems Engineering, Spring 2016
- Korada, J. Directed Individual/Independent study at US veterans Initiative. *Fixed Reconciliation of IT assets of US veterans Initiative and its regional offices* Daniel J. Epstein Department of Industrial and Systems Engineering, Fall 2016
- Xu, T. Directed Individual/Independent study at IMAX Corp., Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2016
- Chithara, R. Directed Individual/Independent study at Camino Real Foods Inc., Daniel J. Epstein Department of Industrial and Systems Engineering, Summer 2015

UNIVERSITY SERVICE

- Daniel J. Epstein Department of Industrial & Systems Engineering, USC
  - MS Analytics Program Director
    - Raised the annual number of new enrolled students from 100 to 300.
    - 60 new Certified (with delayed enrollment for 2022) students in 20213
    - Design of new MS program
    - Developed new MS courses (Business Intelligence, Financial Analytics, Data Analytics Consulting).
    - Webminar for new MS Analytics students, 2020, 2021,2022,2023,2024
  - 4-unit MS Analytics New program design, Committee Chair
  - MS Operations Research Engineering (ORE) Program Advisor
  - 4-unit MS ORE New program design, Committee Chair
  - ISE Department Advisory Board Meeting, November 8, 2018, MS Analytics presentation on current status and modifications.
  - ISE Department MS Awards, 2018, 2023, 2024
  - MS Students Advisement Session, Fall 2018
  - DJE Department of Industrial and Systems Engineering, Faculty Retreat, September 2018.
    A Discussion on USC MS Analytics newly graduate program

  - Faculaty Annual Merit Review Committee, Member, 2017, 2018, 2024
  - ISE Undergraduate program Committee, Member, 2017-2022
- Viterbi School of Engineering, USC
  - Faculty Marshal at the 2018 Viterbi Undergraduate Commencement Ceremony
  - Trojan Family Weekend, Breakfast Reception, Oct 12, 2018
  - Explore USC Reception, February 27, 2017, 2018
  - Viterbi Undergraduate Commencement Ceremony, Faculty Marshal, May 12, 2017
  - Explore USC Reception, February 22, 2016
  - On Campus Admitted Students Reception, OSCAR Meeting, April 24, 2016
  - New Students Luau event, August 20, 2016.

PROFESSION SERVICE

- Quality Engineering Editorial Board Member
- Communications in Statistics - Simulation and Computation - Reviewer
- Computers and Industrial Engineering - Reviewer
- Computational Statistics and Data Analysis - Reviewer
- Book Proposal Reviewer, CRC Press

PROFESSIONAL SOCIETIES

- The European Network of Business and Industrial Statistics (ENBIS)
- Los Angeles *R* Users Group