Jay M. Patel

jay@jaympatel.com | jaypatel87@gmail.com
LINKEDIN: https://www.linkedin.com/in/jay-m-patel-engg
HOMEPAGE: http://www.jaympatel.com
Click here for latest PDF version

PROFILE:

- Data scientist with over five year's experience in data analytics, machine learning, statistics and text mining.
- I have coauthored 1 book, 10 papers, 26 conference presentations and am passionate about explaining data science to non-technical business audiences.
- Frequent speaker at data science events hosted by Federal community of practice (CoP) as part of DigitalGov initiative.

TECHNICAL SKILLS:

- **Machine Learning:** classification and regression (linear, logistic, support vector machine, random forest, convolutional neural network (CNN/ConvNet), recurrent neural networks), cluster analysis, feature engineering. Analyzing unstructured data using Natural-language processing (content and knowledge based recommender systems).
- **Statistical Methods:** hypothesis testing (ANOVA, t-test) and confidence intervals, correlation (bi-variate, partial, distances), time series, principal component analysis and dimensionality reduction.
- **Software and Programming Languages:** Python (scikit-learn, keras, pandas, matplotlib, numpy, scipy, NLTK, spaCy), R (shiny, knitr, ggplot2, tidyverse, caret), SQL (MySQL, SQLite, MongoDB), Apache Spark (MLib), Apache Hadoop, Weka, Eclipse RCP/Java, KNIME, SPSS, Stata, Microsoft Excel.

PROFESSIONAL EXPERIENCE:

- INDEPENDENT CONSULTANT, MUMBAI, INDIA (05/2018 - Present)
  Data Scientist
  - Provide continued support, bug fixes and feature addition to the data analytics python modules and R packages (HTdescR and HTqsarR) developed by me as part of the ongoing projects at US Environmental Protection Agency.
  - Work with stakeholders in development on best standard practices on regression and statistical modeling by being an official voting member on ASTM Committee E11 on Quality and Statistics.
- US ENVIRONMENTAL PROTECTION AGENCY, ATHENS, GA, USA (12/2015 - 05/2018)
  Office of Research and Development (ORD)
  Data Scientist (ORISE Fellow)
  - Fulfilled all the data science duties for US EPA's and US FDA's joint Tox21 program and chemistry safety for sustainability (CSS) program.
  - Compiled and curated data from various sources and used that to develop machine learning based classification and regression models.

- US ENVIRONMENTAL PROTECTION AGENCY, ATHENS, GA, USA (05/2013 - 11/2015)
  Data Scientist (Contract)
  - Directed development of a predictive machine learning based models (more info) as part of a contract valued at over $150,000 with US federal government (Order no. EP13W000134 and EP14W000201, DBA Patel, Jay).
  - Led a team for development of content based recommender system as part of decision analytics dashboard to generate regulatory intelligence insights by using web scraping plus Natural-language processing based model on unstructured data in HTML and pdf format.
- THE UNIVERSITY OF GEORGIA (08/2010 - 05/2013)
  Franklin College of Arts and Sciences
  Research Associate
  - Designed and applied a virtual screening workflow based on machine learning classification model to identify high activity enzyme mutations and validated it experimentally using site saturation mutagenesis.
  - In a separate project, developed a partial least square model for predicting solvation energies for a enzyme mutation and experimentally validated it.
  - Project resulted in four peer reviewed papers in top international journals (Impact factor ~10).

EDUCATION:

- THE UNIVERSITY OF GEORGIA, ATHENS, GA, USA
  M.S., Chemistry (05/2013)
- INSTITUTE OF CHEMICAL TECHNOLOGY (FORMERLY UICT/UDCT), MUMBAI, INDIA
  B.Tech, Chemical Engineering (06/2010)

SELECT PUBLICATIONS:

- Stevens, C.T., Patel, J. M., Koopmans, M., Olmstead, J., Hilal, S.M., Pope, N., Weber, E. J. & Wolfe, K. (2018) Demonstration of a consensus approach for the calculation of physicochemical properties required for environmental fate assessments. Chemosphere.194, 94-106.
- Stevens, C.T., Patel, J. M., Jones, W. J. & Weber, E. J. (2017) Prediction of hydrolysis products of organic chemicals under environmental pH conditions. Environ. Sci. Tech., 51(9), 5008-5016.
- Patel J.M., Phillips R.S. (2014) Effects of hydrostatic pressure on stereospecificity of secondary alcohol dehydrogenase from Thermoanaerobacter ethanolicus support the role of solvation in enantiospecificity. ACS Catalysis. 4, 692-694.
- Patel J.M. (2009) Biocatalytic synthesis of atorvastatin intermediates. J. Mol. Catal. B: Enzym.. 61, 123-128.

SELECT PRESENTATIONS:

- Patel, J. M., Stevens, C.T., Weber, E. J. Estimation of hydrolysis rate constants for carbamates. American Chemical Society (ACS) Annual Spring Meeting 2017, San Francisco, CA, April 02 - 06, 2017.
- Weber, E. J., Card, M. Patel, J. M., Stevens, C.T. Cheminformatics applications and physicochemical property calculators: a powerful combination for the encoding of process science. Gordon Research Conference: Water, NH, Holderness, June 26 - July 01, 2016.