

Basel, Switzerland
Mobile: +41795735169
thirupathi.pattipaka7@gmail.com

Thirupathi Pattipaka

github.com/thirupathipattipaka
linkedin.com/in/thirupathi-pattipaka

Skills and tools

- ❖ MySQL, Oracle, Teradata, MongoDB, Hadoop, Neo4j
- ❖ Hive, Pig, Spark, MLib
- ❖ Python (nltk, scikit-learn, graphlab), Tensorflow, Keras, R (dplyr, plyr, rpart, e1071, randomforest), SAS, SQL, Java, C++
- ❖ R shiny, ggplot2, Tableau, Qlikview, Python (Matplotlib, bokeh, seaborn)
- ❖ Git/ GitLab/ GitHub

projects

- Developed a classification algorithm to classify J9999 insurance drugs claims and final random forest algorithm has produced 95% both accuracy and ROC AUC
- Classification of the type of infusion vaccine (induction vs. maintenance) using machine learning algorithms, and Extreme Gradient boosting model with 96.5% accuracy and 0.99 AUC has selected as the final algorithm
- Developing and scoping a project to identify immune related adverse events using EMR, insurance claims, clinical trials data

Professional Experience

F. Hoffman La Roche Ltd: Senior Data Scientist, Basel

July 2015 to present

- Processed large complex real-world data sources such as medical claims/registries/EHR, genomic sequencing and image datasets
- Applying various new methodologies (Cox proportional data with Elastic Net regularization, Bayesian Networks, intelligence models), new ways to combine data (clinical and genomic features) and generating new use cases
- Taking part of company level machine learning/ advanced analytics initiatives both in image and non-image data and industry level initiatives (Genomics England, IMI, Flatiron, FMI)
- Training, mentoring and leading junior data scientists

PAREXEL International: Consultant – Advanced analytics, Stockholm

October 2012 - July 2015

- Training, testing and cross validation of various data mining/machine learning algorithms like logistic regression, SVM, KNN, Naïve Bayes, random forests, boosting, bagging and cluster analysis

Karolinska Institute, Stockholm - Statistical data analyst

June 2011 - October 2012

- Data mining and machine learning, text mining and automation of predictive modeling, sample size calculation and multivariate statistical testing, generation of statistical reports and graphs

Exon Solutions, Hyderabad - Lead Data analyst

August 2005 - August 2009

- Churn analysis, Risk modeling, Multivariate Modeling, develop and implementation of credit risk score cards, data analysis & statistical modeling
- Sentiment Analysis, Opinion/Buzz and Reputation Analysis, Ad hoc credit risk analysis, scenario analysis
- Predictive Modeling, Dynamical Modeling, Game Theory, recommendation and collaborative filtering

Education

Udacity - Artificial intelligence, Machine learning Nanodegree

2017

KTH Royal Institute of Technology, Sweden – Medical Engineering, MSc

July 2011

Acharya Nagarjuna University, India - Statistics, MSc

May 2009

Osmania University, India – Biomedical, BE

June 2005