720+

Machine Learning

Projects

Accommodation & Food, Agriculture, Banking & Insurance
Biotechnological & Life Sciences, Construction & Engineering, Education & Research, Emergency & Relief, Finance, Manufacturing,
Government and Public Works, Healthcare, Media & Publishing
Justice, Law and Regulations, Accounting, Real Estate, Rental & Leasing
Utilities, Wholesale & Retail

Himanshu Ramchandani
M.Tech | Data Science

Credit: https://github.com/ashishpatel26/Real-time-ML-Project

Accommodation & Food

Food

- RobotChef - Refining recipes based on user reviews.
- Food Amenities - Predicting the demand for food amenities using neural networks
- Recipe Cuisine and Rating - Predict the rating and type of cuisine from a list of ingredients.
- Food Classification - Classification using Keras.
- Image to Recipe - Translate an image to a recipe using deep learning.
- Calorie Estimation - Estimate calories from photos of food.
- Fine Food Reviews - Sentiment analysis on Amazon Fine Food Reviews.

Restaurant

- Restaurant Violation - Food inspection violation forecasting.
- Restaurant Success - Predict whether a restaurant is going to fail.
- Predict Michelin - Predict the likelihood that restaurant is a Michelin restaurant.
- Restaurant Inspection - An inspection analysis to see if cleanliness is related to rating.
- Sales - Restaurant sales forecasting with LTSM.
- Visitor Forecasting - Reservation and visitation number prediction.
- Restaurant Profit - Restaurant regression analysis.
- Competition - Restaurant competitiveness analysis.
- Business Analysis - Restaurant business analysis project.
- Location Recommendation - Restaurant location recommendation tool and analysis.
- Closure, Rating and Recommendation - Three prediction tasks using Yelp data.
- Anti-recommender - Find restaurants you don't want to attend.
- Menu Analysis - Deeper analysis of restaurants through their menus.
- Menu Recommendation - NLP to recommend restaurants with similar menus.
- Food Price - Predict food cost.

- **Automated Restaurant Report** - Automated machine learning company report.
- **Peer-to-Peer Housing** - The effect of peer to peer rentals on housing.
- **Roommate Recommendation** - A system for students seeking roommates.
- **Room Allocation** - Room allocation process.
- **Dynamic Pricing** - Hotel dynamic pricing calculations.
- **Hotel Similarity** - Compare brands that directly compete
- **Hotel Reviews** - Cluster hotel reviews.
- **Predict Prices** - Predict hotel room rates.
- **Hotels vs Airbnb** - Comparing the two approaches.
- **Hotel Improvement** - Analyse reviews to suggest hotel improvements.
- **Orders** - Order cancellation prediction for hotels.
- **Fake Reviews** - Identify whether reviews are fake/spam.
- **Reverse Image Lodging** - Find your preferred lodging by uploading an image.

Accounting

Machine Learning

- **Chart of Account Prediction** - Using labeled data to suggest the account name for every transaction.
- **Accounting Anomalies** - Using deep-learning frameworks to identify accounting anomalies.
- **Financial Statement Anomalies** - Detecting anomalies before filing, using R.
- **Useful Life Prediction (FirmAI)** - Predict the useful life of assets using sensor observations and feature engineering.
- **AI Applied to XBRL** - Standardized representation of XBRL into AI and Machine learning.

Analytics

- **Forensic Accounting** - Collection of case studies on forensic accounting using data analysis. On the lookout for more data to practise forensic accounting, *please get in touch*
- **General Ledger (FirmAI)** - Data processing over a general ledger as exported through an accounting system.
- **Bullet Graph (FirmAI)** - Bullet graph visualisation helpful for tracking sales, commission and other performance.
- **Aged Debtors (FirmAI)** - Example analysis to invetigate aged debtors.

- **Automated FS XBRL** - XML Language, however, possibly port analysis into Python.

Textual Analysis

- **Financial Sentiment Analysis** - Sentiment, distance and proportion analysis for trading signals.
- **Extensive NLP** - Comprehensive NLP techniques for accounting research.

Data, Parsing and APIs

- **EDGAR** - A walk-through in how to obtain EDGAR data.
- **IRS** - Acessing and parsing IRS filings.
- **Financial Corporate** - Rutgers corporate financial datasets.
- **Non-financial Corporate** - Rutgers non-financial corporate dataset.
- **PDF Parsing** - Extracting useful data from PDF documents.
- **PDF Tabel to Excel** - How to output an excel file from a PDF.

Research And Articles

- **Understanding Accounting Analytics** - An article that tackles the importance of accounting analytics.
- **VLFeat** - VLFeat is an open and portable library of computer vision algorithms, which has Matlab toolbox.

Websites

- **Rutgers Raw** - Good digital accounting research from Rutgers.

Courses

- **Computer Augmented Accounting** - A video series from Rutgers University looking at the use of computation to improve accounting.
- **Accounting in a Digital Era** - Another series by Rutgers investigating the effects the digital age will have on accounting.

Agriculture

Economics

- **Prices** - Agricultural price prediction.

- Prices 2 - Agricultural price prediction.
- Yield - Agricultural analysis looking at crop yields in Ukraine.
- Recovery - Strategic land use for agriculture and ecosystem recovery
- MPR - Mandatory Price Reporting data from the USDA's Agricultural Marketing Service.

Development

- Segmentation - Agricultural field parcel segmentation using satellite images.
- Water Table - Predicting water table depth in agricultural areas.
- Assistant - Notebooks from agricultural assistant.
- Eco-evolutionary - Eco-evolutionary dynamics.
- Diseases - Identification of crop diseases and pests using Deep Learning framework from the images.
- Irrigation and Pest Prediction - Analyse irrigation and predict pest likelihood.

Banking & Insurance

Consumer Finance

- Loan Acceptance - Classification and time-series analysis for loan acceptance.
- Predict Loan Repayment - Predict whether a loan will be repaid using automated feature engineering.
- Loan Eligibility Ranking - System to help the banks check if a customer is eligible for a given loan.
- Home Credit Default (FirmAI) - Predict home credit default.
- Mortgage Analytics - Extensive mortgage loan analytics.
- Credit Approval - A system for credit card approval.
- Loan Risk - Predictive model to help to reduce charge-offs and losses of loans.
- Amortisation Schedule (FirmAI) - Simple amortisation schedule in python for personal use.

Management and Operation

- Credit Card - Estimate the CLV of credit card customers.
- Survival Analysis - Perform a survival analysis of customers.

- Next Transaction - Deep learning model to predict the transaction amount and days to next transaction.
- Credit Card Churn - Predicting credit card customer churn.
- Bank of England Minutes - Textual analysis over bank minutes.
- CEO - Analysis of CEO compensation.

Valuation

- Zillow Prediction - Zillow valuation prediction as performed on Kaggle.
- Real Estate - Predicting real estate prices from the urban environment.
- Used Car - Used vehicle price prediction.

Fraud

- XGBoost - Fraud Detection by tuning XGBoost hyper-parameters with Simulated Annealing
- Fraud Detection Loan in R - Fraud detection in bank loans.
- AML Finance Due Diligence - Search news articles to do finance AML DD.
- Credit Card Fraud - Detecting credit card fraud.

Insurance and Risk

- Car Damage Detective - Assessing car damage with convolution neural networks for a personal auto *claims.*
- Medical Insurance Claims - Predicting medical insurance claims.
- Anomaly
- Claim Denial - Predicting insurance claim denial
- Claim Fraud - Predictive models to determine which automobile claims are fraudulent.
- Claims Anomalies - Anomaly detection system for medical insurance claims data.
- Actuarial Sciences (R) - A range of actuarial tools in R.
- Bank Failure - Predicting bank failure.
- Risk Management - Finance risk engagement course resources.
- VaR GaN - Estimate Value-at-Risk for market risk management using Keras and TensorFlow.
- Compliance - Bank Grievance Compliance Management.
- Stress Testing - ECB stress testing.
- Stress Testing Techniques - A notebook with various stress testing exercises.

- Reverse Stress Test - Given a portfolio and a predefined loss size, determine which factors stress (scenarios) would lead to that loss
- BoE stress test- Stress test results and plotting.
- Recovery - Recovery of money owed.
- Quality Control - Quality control for banking using LDA

Physical

- Bank Note Fraud Detection - Bank Note Authentication Using DNN Tensorflow Classifier and RandomForest.
- ATM Surveillance - ATM Surveillance in banks use case.

Biotechnological & Life Sciences

General

- Programming - Python Programming for Biologists
- Introduction DL - A Primer on Deep Learning in Genomics
- Pose - Estimating animal poses using DL.
- Privacy - Privacy preserving NNs for clinical data sharing.
- Population Genetics - DL for population genetic inference.
- Bioinformatics Course - Course materials for Computational *Biology*and Bioinformatics
- Applied Stats - Applied Statistics for High-Throughput *Biology*
- Scripts - Python scripts for biologists.
- Molecular NN - A mini-framework to build and train neural networks for molecular *biology*.
- Systems Biology Simulations - Systems *biology* practical on writing simulators with F# and Z3
- Cell Movement - LSTM to predict biological cell movement.
- Deepchem - Democratizing Deep-Learning for Drug Discovery, Quantum Chemistry, Materials Science and Biology

Sequencing

- DNA, RNA and Protein Sequencing - Anew representation for biological sequences using DL.
- CNN Sequencing - A toolbox for learning motifs from DNA/RNA sequence data using convolutional neural networks
- NLP Sequencing - Language transfer learning model for genomics

Chemoinformatics and drug discovery

- Novel Molecules - A convolutional net that can learn features.
- Automating Chemical Design - Generate new molecules for efficient exploration.
- GAN drug Discovery - A method that combines generative models with reinforcement learning.
- RL - generating compounds predicted to be active against a biological target.
- One-shot learning - Python library that aims to make the use of machine-learning in drug discovery straightforward and convenient.

Genomics

- Jupyter Genomics - Collection of computation biology and bioinformatics notebooks.
- Variant calling - Correctly identify variations from the reference genome in an individual's DNA.
- Gene Expression Graphs - Using convolutions on an image.
- Autoencoding Expression - Extracting relevant patterns from large sets of gene expression data
- Gene Expression Inference - Predict the expression of specified target genes from a panel of about 1,000 pre-selected “landmark genes”.
- Plant Genomics - Presentation and example material for *Plant* and Pathogen Genomics

Life-sciences

- Plants Disease - App that detects diseases in *plants* using a deep learning model.
- Leaf Identification - Identification of *plants* through *plant* leaves on the basis of their shape, color and texture.
- Crop Analysis - An imaging library to detect and track future position of ears on maize *plants*
- Seedlings - *Plant* Seedlings Classification from kaggle competition
- [Plant Stress](http://An ontology containing plant stresses; biotic and abiotic.) - An ontology containing *plant* stresses; biotic and abiotic.
- Animal Hierarchy - Package for calculating *animal* dominance hierarchies.
- Animal Identification - Deep learning for animal identification.
- Species - Big Data analysis of different species of *animals*
- Animal Vocalisations - A generative network for animal vocalizations
- Evolutionary - Evolution Strategies Tool

- Glaciers - Educational material about glaciers.

Construction & Engineering

Construction

- DL Architecture - Deep learning classifier and image generator for building architecture.
- Construction Materials - A course on construction materials.
- Bad Actor Risk Model - Risk model to improve construction related building safety
- Inspectors - Determine the assigned inspections.
- Corrupt Social Interactions - Uncover potential corrupt social interactions between an industry member and the staff at the DOB
- Risk Construction - Identify high risk construction.
- Facade Risk - A risk model to predict unsafe facades.
- Staff Levels - Predicting staff levels for front line workers.
- Injuries - Building related injuries topic modelling.
- Building Violations - Predictive analysis of building violations.
- Productivity - Productivity analysis and inspection with Tableau.

Engineering:

- Structural Analysis - 2D Structural Analysis in Python.
- Structural Engineering - Structural engineering modules.
- Nusa - Structural analysis using the finite element method.
- StructPy - Structural Analysis Library for Python based on the direct stiffness method
- Aileron - Structural analysis of the aileron of a Boeing 737
- Vibration - Educational vibration programs.
- Civil - Collection of civil engineering tools in FreeCAD
- GEstimator - Simple civil estimation software
- Fatpack - Functions and classes for fatigue analysis of data series.
- Pysteel - Automated design of different steel structure
- Structural Uncertainty - Quantifying structural uncertainty with deep learning.
- Pymech - A Python module for mechanical engineers
- Aerospace Engineering - Astrodynamics and Statistics
- Interactive Quantum Chemistry - Combining Psi4 and Numpy for education and development.
- Chemical and Process Engineering - Various resources.

- PyTherm - Applied Thermodynamics
- Aerogami - Aerodynamics using planes.
- Electro geophysics - Interactive applications for electromagnetics in geophysics
- Graph Signal - Graph signal processing tutorial.
- Mechanical Vibrations - Mechanical Vibrations at the Univsersity of Louisiana.
- Process Dynamics - Process Dynamics and Control
- Battery Life Cycle - Data driven prediction of batter life cycle.
- Wind Energy - Python for wind energy
- Energy Use - Standard methods for calculating normalized metered energy consumption
- Nuclear Radiation - How people are affected by radiations emitted by nuclear power plants

Material Science

- Python Materials Genomics - Robust material analysis code used in a well-established project.
- Materials Mining - Scripts for simulations and analysis of materials.
- Emmet - Build databases of material properties.
- Megnet - Graph networks as a ML framework for Molecules and Crystals
- Atomate - Pre-built workflows for computational material science.
- Bylaws Compliance - Predicting property fines.
- Asphalt Binder - Construction materials, free energy and chemical composition of asphalt binder.
- Steel - Optimisation of steel.
- Awesome Materials Informatics - Curated list of known efforts in materials informatics.

Economics

General

- Trading Economics API - Information for 196 countries.
- Development Economics - Development microeconomics are written mostly as interactive jupyter notebooks
- Applied Econ & Fin - Applied Computational Economics and Finance
- Macroeconomics - Topics in macroeconomics with notebook examples.

Machine Learning

- **EconML** - Automated Learning and Intelligence for Causation and *Economics.*
- **Auctions** - Optimal auctions using deep learning.

Computational

- **Quant Econ** - Quantitative economics course by NYU
- **Computational** - Computational methods in economics.
- **Computational 2** - Small course in computational economics.
- **Econometric Theory** - Notebooks of A Primer on Econometric theory.

Education & Research

Student

- **Student Performance** - Mining student performance using machine learning.
- **Student Performance 2** - Student exam performance.
- **Student Performance 3** - Student achievement in secondary education.
- **Student Performance 4** - Students Performance Evaluation using Feature Engineering
- **Student Intervention** - Building a student intervention system.
- **Student Enrolment** - Student enrolment and performance analysis.
- **Academic Performance** - Explore the demographic and family features that have an impact a student's academic performance.
- **Grade Analysis** - Student achievement analysis.

School

- **School Choice** - Data analysis for education's school choice.
- **School Budgets and Priorities** - Helping the school board and mayor make strategic decisions regarding future school budgets and priorities
- **School Performance** - Data analysis practice using data from data.utah.gov on school performance.
- **School Performance 2** - Using pandas to analyze school and student performance within a district
- **School Performance 3** - Philadelphia School Performance
- **School Performance 4** - NJ School Performance
- **School Closure** - Identify schools at risk for closure by performance and other characteristics.
- **School Budgets** - Tools and techniques for school budgeting.

- School Budgets - Same as a above, datacamp.
- PyCity - School analysis.
- PyCity 2 - School budget vs school results.
- Budget NLP - NLP classification for budget resources.
- Budget NLP 2 - Further classification exercise.
- Budget NLP 3 - Budget classification.
- Survey Analysis - Education survey analysis.

Emergency & Police

Preventative and Reactive

- Emergency Mapping - Detection of destroyed houses in California
- Emergency Room - Supporting em*ergency r*oom decision making
- Emergency Readmission - Adjusted Risk of *Emergency* Readmission.
- Forest Fire - Forest fire detection through UAV imagery using CNNs
- Emergency Response - Emergency response analysis.
- Emergency Transportation - Transportation prompt on *emergency* services
- Emergency Dispatch - Reducing response times with predictive modeling, optimization, and automation
- Emergency Calls - Emergency calls analysis project.
- Calls Data Analysis - 911 data analysis.
- Emergency Response - Chemical factory RL.

Crime

- Crime Classification - Times analysis of serious assaults misclassified by LAPD.
- Article Tagging - Natural Language Processing of Chicago news article
- Crime Analysis - Association Rule Mining from Spatial Data for *Crime* Analysis
- Chicago Crimes - Exploring public Chicago *crimes* data set in Python
- Graph Analytics - The Hague Crimes.
- Crime Prediction - *Crime* classification, analysis & prediction in Indore city.
- Crime Prediction - Developed predictive models for *crime* rate.
- Crime Review - Crime review data analysis.
- Crime Trends - The *Crime* Trends Analysis Tool analyses *crime* trends and surfaces problematic *crime* conditions
- Crime Analytics - Analysis of *crime* data in Seattle and San Francisco.

Ambulance:

- Ambulance Analysis - An investigation of Local Government Area ambulance time variation in Victoria.
- Site Location - Ambulance site locations.
- Dispatching - Applying game theory and discrete event simulation to find optimal solution for ambulance dispatching
- Ambulance Allocation - Time series analysis of ambulance dispatches in the City of San Diego.
- Response Time - An analysis on the improvements of ambulance response time.
- Optimal Routing - Project to find optimal routing of ambulances in Ithaca.
- Crash Analysis - Predicting the probability of accidents on a given segment on a given time.

Disaster Management

- Conflict Prediction - Notebooks on conflict prediction.
- Burglary Prediction - Spatio-Temporal Modelling for burglary prediction.
- Predicting Disease Outbreak - Machine Learning implementation based on multiple classifier algorithm implementations.
- Road accident prediction - Prediction on type of victims on federal road accidents in Brazil.
- Text Mining - Disaster Management using Text mining.
- Twitter and disasters - Try to correctly predict whether tweets that are about disasters.
- Flood Risk - Impact of catastrophic flood events.
- Fire Prediction - We used 4 different algorithms to predict the likelihood of future fires.

Finance

Trading and Investment

- For more see financial-machine-learning
- Deep Portfolio - Deep learning for finance Predict volume of bonds.
- AI Trading - Modern AI trading techniques.
- Corporate Bonds - Predicting the buying and selling volume of the corporate bonds.
- Simulation - Investigating simulations as part of computational finance.
- Industry Clustering - Project to cluster industries according to financial attributes.
- Financial Modeling - HFT trading and implied volatility modeling.

- Trend Following - A futures trend following portfolio investment strategy.
- Financial Statement Sentiment - Extracting sentiment from financial statements using neural networks.
- Applied Corporate Finance - Studies the empirical behaviors in stock market.
- Market Crash Prediction - Predicting market crashes using an LPPL model.
- NLP Finance Papers - Curating quantitative finance papers using machine learning.
- ARIMA-LTSM Hybrid - Hybrid model to predict future price correlation coefficients of two assets
- Basic Investments - Basic investment tools in python.
- Basic Derivatives - Basic forward contracts and hedging.
- Basic Finance - Source code notebooks basic finance applications.
- Advanced Pricing ML - Additional implementation of Advances in Financial Machine Learning (Book)
- Options and Regression - Financial engineering project for option pricing techniques.
- Quant Notebooks - Educational notebooks on quant finance, algorithmic trading and investment strategy.
- Forecasting Challenge - Financial forecasting challenge by G-Research (Hedge Fund)
- XGboost - A trading algorithm using XgBoost
- Research Paper Trading - A strategy implementation based on a paper using Alpaca Markets.
- Various - Options, Allocation, Simulation
- ML & RL NYU - Machine Learning and Reinforcement Learning in Finance.

Data

- Datastream - Datastrem from Thomson Reuters accessible through Python.
- AlphaVantage - API wrapper to simplify the process of acquiring free financial data.
- FSA- A project to transfer SEC Edgar Filings' financial data to custom financial statement analysis models.
- TradeConnector - A layer to connect with market data providers.
- Employee Count SEC Filings
- SEC Parsing
- Open Edgar
- Rating Industries

Healthcare

General

- zEpid - Epidemiology analysis package.
- Python For Epidemiologists - Tutorial to introduce epidemiology analysis in Python.
- Prescription Compliance - An analysis of prescription and medical compliance
- Respiratory Disease - Tracking respiratory diseases in Olympic athletes
- Bubonic Plague - Bubonic plague and SIR model.

Justics, Law & Regulations

Tools

- LexPredict - Software package and library.
- AI Para-legal - Lobe is the world's first AI paralegal.
- Legal Entity Detection - NER For Legal Documents.
- Legal Case Summarisation - Implementation of different summarisation algorithms applied to legal case judgements.
- Legal Documents Google Scholar - Using Google scholar to extract cases programatically.
- Chat Bot - Chat-bot and email notifications.
- Congress API - ProPublica congress API access.
- Data Generator GDPR - Dummy data generator for GDPR compliance

Policy and Regulatory

- GDPR scores - Predicting GDPR Scores for Legal Documents.
- Driving Factors FINRA - Identify the driving factors that influence the FINRA arbitration decisions.
- Securities Bias Correction - Bias-Corrected Estimation of Price Impact in Securities Litigation.
- Public Firm to Legal Decision - Embed public firms based on their reaction to legal decisions.
- Night Life Regulation - Australian nightlife and its regulation and policing
- Comments - Public comments on government regulations.
- Clustering - Clustering Canadian regulations.
- Environment - Regulation of Energy and the Environment

- **Risk** - Systematic risk of various financial regulations.
- **FINRA Compliance** - Topic modelling on compliance.

Judicial Applied

- **Supreme Court Prediction** - Predicting the ideological direction of Supreme Court decisions: ensemble vs. unified case-based model.
- **Supreme Court Topic Modeling** - Multiple steps necessary to implement topic modeling on supreme court decisions.
- **Judge Opinion** - Using text mining and machine learning to analyze judges' opinions for a particular concern.
- **ML Law Matching** - A machine learning law match maker.
- **Bert Multi-label Classification** - Fine Grained Sentiment Analysis from AI.
- **Some Computational AI Course** - Video series Law MIT.

Manufacturing

General

- **Green Manufacturing** - Mercedes-Benz Greener *Manufacturing* competition on Kaggle.
- **Semiconductor Manufacturing** - Semicondutor *manufacturing* process line data analysis.
- **Smart Manufacturing** - Shared work of a modelling Methodology.
- **Bosch Manufacturing** - Bosch manufacturing project, Kaggle.

Maintenance

- **Predictive Maintenance 1** - Predict remaining useful life of aircraft engines
- **Predictive Maintenance 2** - Time-To-Failure (TTF) or Remaining Useful Life (RUL)
- **Manufacturing Maintenance** - Simulation of maintenance in *manufacturing* systems.

Failure

- **Predictive Analytics** - Method for Predicting failures in Equipment using Sensor data.
- **Detecting Defects** - Anomaly detection for defective semiconductors
- **Defect Detection** - Smart defect detection for pill manufacturing.
- **Manufacturing Failures** - Reducing manufacturing failures.

- **Manufacturing Anomalies - Intelligent anomaly detection for *manufacturing* line.**

Quality

- **Quality Control - Bosh failure of quality control.**
- **Manufacturing Quality - Intelligent *Manufacturing* Quality Forecast**
- **Auto Manufacturing - Regression Case Study Project on *Manufacturing* Auction Sale Data.**

Media & Publishing

Marketing

- **Video Popularity - HIP model for predicting the popularity of videos.**
- **YouTube transcriber - Automatically transcribe YouTube videos.**
- **Marketing Analytics - Marketing analytics case studies.**
- **Algorithmic Marketing - Models from Introduction to Algorithmic Marketing book**
- **Marketing Scripts - Marketing data science applications.**
- **Social Mining - Mining the social web.**

Miscellaneous

Art

- **Painting Forensics - Analysing paintings to find out their year of creation.**

Tourism

- **Flickr - Metadata mining tool for tourism research.**
- **Fashion - A clothing retrieval and visual recommendation model for fashion images**

Physics

General

- **Gamma-hadron Reconstruction - Tools used in Gamma-ray ground based astronomy.**
- **Curriculum - Newtonian notebooks.**

- Interaction Networks - Interaction Networks for Learning about Objects, Relations and *Physics.*
- Particle Physics - Training, generation, and analysis code for learning Particle *Physics*
- Computational Physics - A computational physics repository.
- Medical Physics - Useful python for medical physics.
- Medical Physics 2 - A common, core Python package for Medical *Physics*
- Flow Physics - Flow *Physics* and Aeroacoustics Toolbox with Python

Machine Learning

- Physics ML and Stats - Machine learning and statistics for physicists
- High Energy - Machine Learning for High Energy *Physics*.
- High Energy GAN - Generative Adversarial Networks for High Energy *Physics.*
- Neural Networks - P*hysics* meets neural networks

Government and Public Works

Social Policies

- Triage - General Purpose Risk Modeling and Prediction Toolkit for Policy and Social Good Problems.
- World Bank Poverty I - A comparative assessment of machine learning classification algorithms applied to poverty prediction.
- World Bank Poverty II - Repository for the World Bank Pover-t Test Competition Solution Overseas Company Land Ownership .
- Overseas Company Land Ownership - Identifying foreign ownership in the UK.
- CFPB - Consumer Finances Protection Bureau complaints analysis.
- Cannabis Legalisation Effect - Effects of cannabis legalization on crime.
- Public Credit Card - Identification of potential fraud for council credit cards. Data
- Recidivism Prediction - Transparency and audibility to recidivism risk assessment
- Household Poverty - Predict poverty in households in Costa Rica.
- NLP Public Policy - An example of an NLP use-case in public policy.
- World Food Production - Comparing Top food and feed Producers around the globe.
- Tax Inequality - Data project around taxation and inequality in Basel Stadt.

- **Sheriff Compliance - Compliance to ICE requests.**
- **Apps Detection - Suspicious app detection for kids.**
- **Social Assistance - Trending information on social assistance**
- **Computational Social Science - Social data science summer school course.**
- **Liquor and Crime - Effect of liquor licenses issued on the crime rate.**
- **Animal Placement Kennels - Optimising animal placement in shelters.**
- **Staffing Wall - Independent exploration project on U.S. Mexican Border wall**
- **Worker Fatalities - Worker Fatalities and Catastrophes Map from OSHA data**

Charities

- **Census Data API - Pull variables from the 5-year American Community Survey.**
- **Philantropic Giving - Work done by numerous DataKind volunteers on harnessing Form 990 data**
- **Charity Recommender - NYC *Charity* Collaborative Recommender System on an Implicit DataSet.**
- **Donor Identification - A machine learning project in which we need to find donors for *charity.***
- **US Charities - Charity exploration and machine learning.**
- **Charity Effectiveness**
    - **Scraping online data about *charities* to understand effectiveness**

Election Analysis

- **Election Analysis - Election Analysis and Prediction Models**
- **American Election Causal - Using ANES data with causal inference models.**
- **Campaign Finance and Election Results - Investigating the relation between campaign finance and subsequent election results.**
- **Voting System - Proportional representation voting methods.**
- **President Vote - Vote by income level analysis..**

Politics

- **Congressional politics - House and senate congressional partisanship.**
- **Politico - A platform for profiling public figures in Brazilian *politics.***
- **Bots - Tools and algorithms to analyze Paraguayan Tweets in times of election**
- **Gerrymander tests - Lots of metrics for quantifying gerrymandering.**
- **Sentiment - Analyse newspapers with respect to their *political* conviction using entity sentiments of party representatives.**

- DL Politics - Prediction of Spanish *Political* Affinity with Deep Neural Nets: Socialist vs People's Party
- PAC Money - Effects of PAC money on US *politics*.
- Power Networks - Constructing a watchdog for Indian corporate and *political* networks
- Elite - Political elite in the US.
- Debate Analysis - Program to analyze *political* debates.
- Political Affiliation - Political affiliation prediction using twitter metadata.
- Political Ads - Investigation into Facebook *Political* Ads and Targeting
- Political Identity - Multi-axial *political* model.
- YT Politics - Mapping *Politics* on YouTube
- Political Ideology - Unsupervised learning of *political* ideology by word vector projections

Real Estate, Rental & Leasing

Real Estate

- Finding Donuts - Finding real estate opportunities by predicting transforming neighbourhoods.
- Neighbourhood - Predicting real estate prices from the urban environment.
- Real Estate Classification - Classifying the type of property given Real Estate, satellite and Street view Images
- Recommender - This tools aims to recommend a user the top 5 real estate properties that matches their search.
- House Price - Predicting *house* prices using Linear Regression and GBR
- House Price Portland - Predict housing prices in Portland.
- Zillow Prediction - Zillow valuation prediction as performed on Kaggle.
- Real Estate - Predicting real estate prices from the urban environment.

Rental & Leasing

- Analysing Rentals - Analyzing and visualizing rental listings data.
- Interest Prediction - Predict people interest in renting specific NYC apartments.
- Housing Uni vs Non-Uni - The effect on university lodging after the GFC.
- Predict Household Poverty - Predict the poverty of households in Costa Rica using automated feature engineering.
- Airbnb public analytics competition: - Now strategic management.

Utilities

Electricity

- Electricity Price - Electricity price comparison Singapore.
- Electricity-Coal Correlation - Determining the correlation between state electricity rates and coal generation over the past decade.
- Electricity Capacity - A Los Angeles Times analysis of California's costly power glut.
- Electricity Systems - Optimal Wind+Hydrogen+Other+Battery+Solar (WHOBS) *electricity* systems for European countries.
- Load Disaggregation - Smart meter load disaggregation with Hidden Markov Models
- Price Forecasting - Forecasting Day-Ahead *electricity* prices in the German bidding zone with deep neural networks.
- Carbon Index - Calculation of *electricity* CO_2 intensity at national, state, and NERC regions from 2001-present.
- Demand Forecasting - *Electricity* demand forecasting for Austin.
- Electricity Consumption - Estimating *Electricity* Consumption from Household Surveys
- Household power consumption - Individual household power consumption LSTM.
- Electricity French Distribution - An analysis of electricity data provided by the French Distribution Network (RTE)
- Renewable Power Plants - Time series of cumulated installed capacity.
- Wind Farm Flow - A repository of wind plant flow models connected to FUSED-Wind.
- Power Plant - The dataset contains 9568 data points collected from a Combined Cycle Power Plant over 6 years (2006-2011).

Coal, Oil & Gas

- Coal Phase Out - Generation adequacy issues with Germany’s coal phaseout.
- Coal Prediction - Predicting coal production.
- Oil & Gas - Oil & *Natural Gas* price prediction using ARIMA & Neural Networks
- Gas Formula - Calculating potential economic effect of price indexation formula.
- Demand Prediction - Natural gas demand prediction.
- Consumption Forecasting - Natural gas consumption forecasting.

- **Gas Trade - World Model for *Natural Gas* Trade.**

Water & Pollution

- **Safe Water - Predict health-based drinking water violations in the United States.**
- **Hydrology Data - A suite of convenience functions for exploring water data in Python.**
- **Water Observatory - Monitoring water levels of lakes and reservoirs using satellite imagery.**
- **Water Pipelines - Using machine learning to find water pipelines in aerial images.**
- **Water Modelling - Australian Water Resource Assessment (AWRA) Community Modelling System.**
- **Drought Restrictions - A Los Angeles Times analysis of water usage after the state eased drought restrictions**
- **Flood Prediction - Applying LSTM on river water level data**
- **Sewage Overflow - Insights into the sanitary sewage overflow (SSO).**
- **Water Accounting - Assembles water budget data for the US from existing data source**
- **Air Quality Prediction - Predict air quality(aq) in Beijing and London in the next 48 hours.**

Transportation

- **Transdim - Creating accurate and efficient solutions for the spatio-temporal traffic data imputation and prediction tasks.**
- **Transport Recommendation - Context-Aware Multi-Modal Transportation Recommendation**
- **Transport Data - Data and notebooks for Toronto transport.**
- **Transport Demand - Predicting demand for public transportation in Nairobi.**
- **Demand Estimation - Implementation of dynamic origin-destination demand estimation.**
- **Congestion Analysis - Transportation systems analysis**
- **TS Analysis - Time series analysis on transportation data.**
- **Network Graph Subway - Vulnerability analysis for transportation networks.**
- **Transportation Inefficiencies - Quantifying the inefficiencies of Transportation Networks**
- **Train Optimisation - Train schedule optimisation**
- **Traffic Prediction - multi attention recurrent neural networks for time-series (city traffic)**

- **Predict Crashes** - Crash prediction modelling application that leverages multiple data sources
- **AI Supply chain** - Supply chain optimisation system.
- **Transfer Learning Flight Delay** - Using variation encoders in Keras to predict flight delay.
- **Replenishment** - Retail replenishment code for supply chain management.

Wholesale & Retail

Wholesale

- **Customer Analysis** - Wholesale customer analysis.
- **Distribution** - JB wholesale distribution analysis.
- **Clustering** - Unsupervised learning techniques are applied on product spending data collected for customers
- **Market Basket Analysis** - Instacart public dataset to report which products are often shopped together.

Retail

- **Retail Analysis** - Studying Online *Retail* Dataset and getting insights from it.
- **Online Insights** - Analyzing the Online Transactions in UK
- **Retail Use-case** - Notebooks & Data for CyberShop *Retail* Use Case
- **Dwell Time** - Customer dwell time and other analysis.
- **Retail Cohort** - Cohort analysis.

Credit: https://github.com/ashishpatel26/Real-time-ML-Project

Data Science ML Full Stack Roadmap
https://github.com/hemansnation/Data-Science-ML-Full-Stack-2022

Join Telegram for Data Science ML AI Resources:
https://t.me/+sREuRiFssMo4YWJl

Connect with me on these platforms:
LinkedIn: https://www.linkedin.com/in/hemansnation/

Twitter: https://twitter.com/hemansnation

GitHub: https://github.com/hemansnation

Instagram: https://www.instagram.com/masterdexter.ai/

Are you a professional?
DM for One-on-One sessions for Python, Data Science, Machine Learning, and Data Engineering.
Here: https://bit.ly/3U6zQvQ