

Predicting Rate of Adverse Effects and Identifying driving factors

Revenue Drivers of Pharma – Drug Sales

Clinical Trial Pipelines

1. pre-clinical research
2. CT - 1,2,3
3. FDA review
4. Post market monitoring

Network

- Biotech firms
- Marketing agencies
- Contract Research

Trade

- Export and import markets
- API Sourcing – active pharma ingredient

IP

- Patents – citations, expiry and additional R&D

Investments in pharmaceutical companies crossed \$1 Billion post COVID. Globally, this is set to reach 957 Billion by 2028.

Biotech firms have lots of uncertainty in success of drug development and commercialization. Their research pipelines can be valued by studying clinical trials data.

Harvard study in 2020 revealed more drugs are approved just before holidays and they have a high rate of adverse effects.

Source	Description
AACT	Clinical trials data
Drug Approvals	Date of drug approval
Open Payments	Sponsoring from Pharma to doctors
FAERS	Repo of Adverse events
Drug Utilization	Sales with insurance cover

5 data sources from publicly available FDA data
Text and tabular data

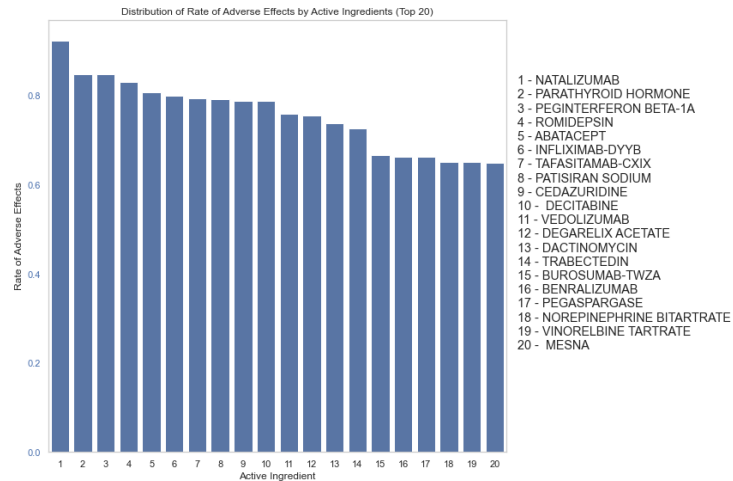
Approval_Month	# of Approvals	# of adverse_affects	Adverse_intensity
1	1.63%	0.24%	0.15
2	1.57%	0.23%	0.15
3	2.33%	0.17%	0.07
4	3.08%	1.02%	0.33
5	3.40%	0.79%	0.23
6	3.90%	1.02%	0.26
7	5.10%	1.01%	0.20
8	6.85%	1.75%	0.26
9	9.13%	7.96%	0.87
10	11.70%	6.32%	0.54
11	18.70%	12.88%	0.69
12	32.62%	66.61%	2.04

Desk clearing nature before holidays and at end of month is causing a lax review of drugs in USA, Europe and China.

Drugs used to treat terminal diseases have a general high rate of adverse effects. Benralizumab used for Asthma is an outlier

Natalizumab is #1 API with the highest rate of adverse effects. Used to treat cancer and multiple sclerosis.

Romidepsin, Trabectedin and Dactinomycin – Treat Cancer
Abatacept and Peginterferon beta-1a – Treat Autoimmune



	Tuning params	Total Fits	Time Taken (s)	Train F1 Score	Test F1 Score
Decision Tree	24	120	24.32	0.9517	0.7718
Random Forest	9	45	113.93	0.9966	0.7995
LightGBM	15	75	11.74	0.9583	0.8005
Logistic Regression	5	25	27.17	0.7769	0.7772
SVM Classifier	8	40	5.37	0.8250	0.7551
KNN Classifier	8	40	7.98	0.7827	0.7843
Voting Classifier	-	-	10.04	0.9875	0.8079

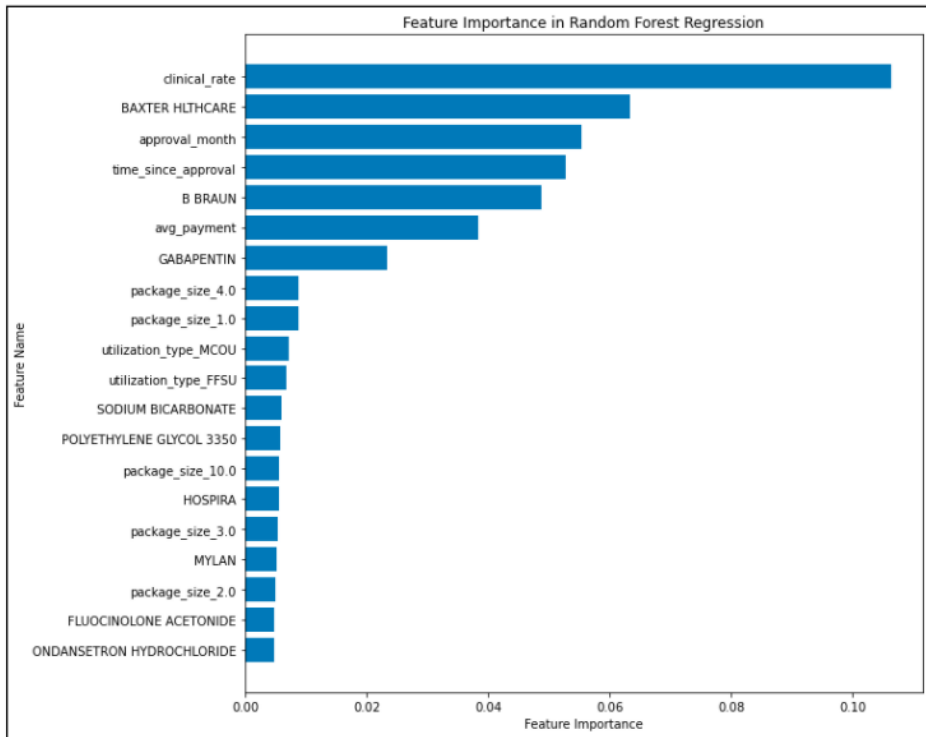
Classification and Regression based Modelling Approaches to predict rate of adverse effects.

Model	Train		Test	
	MSE	R ²	MSE	R ²
Ridge Regression	1.5	0.78	2.5	0.62
Random Forest Regression	0.25	0.96	1.59	0.76

80-20 Train-Test split

Tree based regression models have best performance when we expect data can be split into groups

Conclusion



Adverse effects seen during clinical trials is strongly indicative of adverse effects post approval in the market

Drugs approved in December have the highest rate of adverse effects

Increase in rate of adverse effects till the 2nd year since approval and a notable decrease immediately after

Pharma companies sponsor doctors and the average total payment made is a strong driver for adverse effects

Package sizes also come up an important feature driver

Companies with a high rate of adverse effects have had concerns raised in the media due to their drugs

Biogen Inc had a 40% plunge in its stock price after the death of an Alzheimer's patient due to Adulheim



The Team



Archit Matta



Yosha Singh Tomar



Aditya Koduri



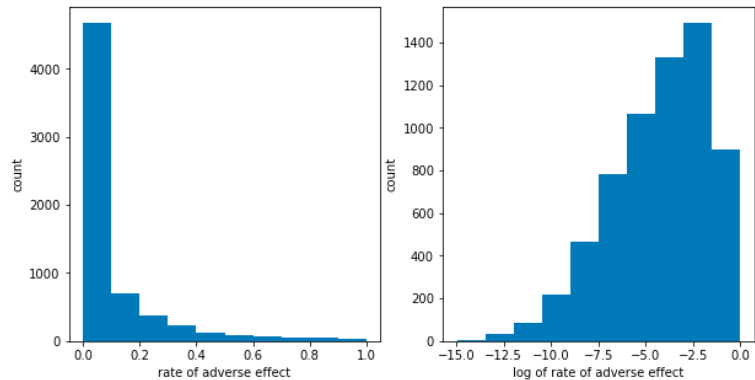
Karunakar Gadireddy



Shivani Modi



Histogram of target variable



SAME NAME, LOW SIMILARITY

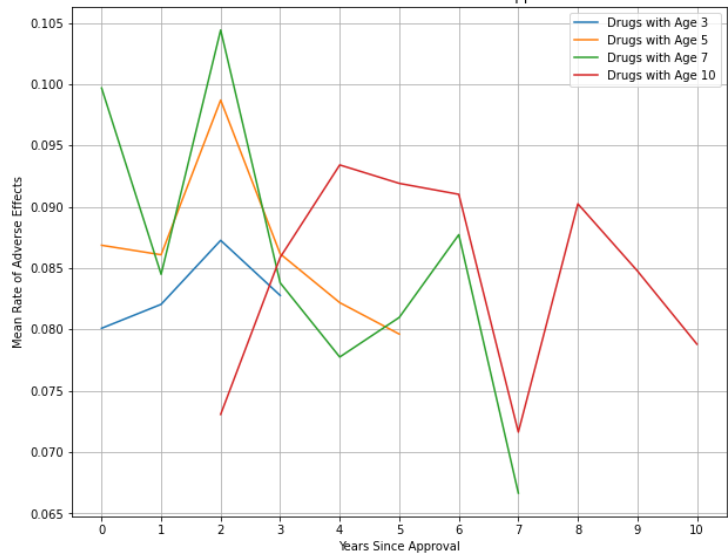
ORGANON	OrganOx Ltd	0.707
BAYER HLTHCARE	Bayer HealthCare Pharmaceuticals Inc.	0.684
ABBVIE ENDOCRINE INC	AbbVie, Inc.	0.618
SANOFI AVENTIS US	Sanofi Pasteur Inc.	0.455
GLAXO GRP LTD	GlaxoSmithKline, LLC.	0.375
OTSUKA PHARM CO LTD	Otsuka Pharmaceutical Development & Commercial...	0.374

DIFFERENT NAME, HIGH SIMILARITY

HOSPIRA	Spiration, Inc.	0.577
PAR STERILE PRODUCTS	Synthes USA Products LLC	0.512
COVIS	Covidien LP	0.483
KING PHARMS LLC	NS Pharma, Inc.	0.323
SOMERSET	Intersect ENT, Inc.	0.319

TF-IDF based polyfuzz matching after exclusion of stop words

Rate of Adverse Effects vs Years Since Approval



Frequent Names in Company Name

