

YASH KALPESH PANCHAL

M. Tech, Industrial and Management Engineering, IIT Kanpur

[linkedin.com/in/yash-panchal-83385bb1/](https://www.linkedin.com/in/yash-panchal-83385bb1/)
ypanchal@iitk.ac.in | +91-8460288427

ACADEMIC QUALIFICATION

YEAR	QUALIFICATION	EDUCATIONAL INSTITUTION	%/CPI
2019-cont.*	M. Tech, Industrial & Management Engineering	Indian Institute of Technology, Kanpur	9.01 CPI
2011-15	B. Tech, Mechanical Engineering	Dharmsinh Desai Institute of Technology, Nadiad	8.18 CPI
2011	Class XII GHSE Board	Shree Narayana Guru Vidhyalaya, Ahmedabad	80.00%
2009	Class X GSE Board	Karnavati Public School, Ahmedabad	86.31%

*upto 2nd sem

WORK EXPERIENCE

Quality Assurance Engineer, Indian Space Research Organization

(Feb'17-June'18)

- **Qualified and approved design**, assembly process, drawings of Packages performing Vibration Tests, Thermo-Vacuum Test, etc.
- Audited **GSAT-6A, GSAT-29** projects, by developing spreadsheet recording detailed history of all packages.

Design Engineer, Indian Space Research Organization

(May'15-May'16)

- Developed Mechanical Structure of **CHANDRAYAAN-II**, Imaging Infrared Spectrometer, Terrain Mapping Camera-2 and Rover Imager
- **Designed**, transportation and handling purpose **container** for TMC-2, maintaining specific range of Humidity, Temperature, O₂-N₂ %, thermal shock.

INTERNSHIP

Data Science Intern, Harvesting India Private Limited

(April'20-June'20)

Crop Classification using Supervised Deep Learning and Unsupervised Model

- Classified crops based on Remote Sensing Data, consisting **Multi-Spectral (5 bands)** Raster image files (.tif) with **spatial resolution of 10m**.
- Used **QGIS software** for resolution setting, mapping, visualization of images.
- Initiated model development by examining **Bathinda District data**, Python Libraries – **GDAL, geopandas, PIL, Sci-Kit Image** for data pre-processing.
- Used **Normalized Difference Vegetation Index (NDVI)** for assessing green vegetation presence over the Raster Image.
- Applied **Logistic Regression, Random Forest, Deep Neural Network** with multi-layers and compared these models on common metric.
- Expanded the idea to **unsupervised model**, by applying, **K-means Clustering, Gaussian Mixture Model (GMM)**.
- **Precision, Recall, F1-Score and Accuracy** – 0.82, 0.91, 0.86, and 0.82, analyzed **DHAR district** with unsupervised K-means, **Multi-Class Classification**.

ACADEMIC PROJECTS

Amazon Fine Food Review Classification – Data Mining and Knowledge Discovery

(Aug'19-Nov'19)

- Classified reviews based on sentimental data, performed data pre-processing by removing duplicate entries, Stopwords, Snowball Stemming.
- Executed Feature Extraction Techniques – **Bag of Words, TF-IDF, Average Word2Vec, TF-IDF Word2Vec**.
- Applied Models – **Logistic Regression, Random Forest, Support Vector Machine, Naïve Bayes, Neural Network, K-Nearest Neighbors**, with k-folds **CV**.
- Used **Accuracy, Precision, Recall and F1-Score** as metrics for comparison yielding, **Logistic Regression with TF-IDF as winner with 0.93 accuracy**.

Predict Movie's Worldwide Box Office Revenue – Data Mining and Knowledge Discovery

(Aug'19-Nov'19)

- Predicted Movie's Revenue Collection with 22 independent variables (IV), performed **Explanatory Data Analysis (EDA) and Feature Engineering (FE)**.
- Models - **OLS Linear Regression, ensemble models - Random Forest, Extra Trees, XGBoost**, with **k-fold CV**, each tuned using BayesSearchCV function.
- Compared models with RMSE metric, which yielded XGBoost as the best model for given EDA, FE and dataset with **RMSE as 2.04**

Multivariate Linear Regression Model to Estimate Student's Performance – Statistical Modeling for Business Analytics

(Jan'20-Feb'20)

- Developed a best fit Multivariate Linear Regression Model for estimating Students performance (marks).
- Feature Engineered variables, Categorical to numerical, used **Recursive Feature Elimination** technique for **Feature Selection based on p-value**.
- Best fit model consisted **8 variables having 95% significance** which resulted in **R² as 0.32 and Adjusted R² as 0.30**
- Qualified model by **Breusch Pagan (BP) Test: Heteroscedasticity, Variation Inflation Factor (VIF)** for Multicollinearity, checked **omitted variable bias**.

Binary Classification of Occupancy of Room – Statistical Modeling for Business Analytics

(April'20-July'20)

- Binary Classification of occupancy of room, based on variables – temperature, Humidity, Light, CO₂ content, Humidity Ratio.
- Applied Non-linear models, **Logit Regression, Probit Regression** for determining the probability of the occupancy.
- Performed Tests – **Variation Inflation Factor (VIF)** for multicollinearity check, **Variable Significance Test @ 95% level**.
- Evaluated models on **ROC curve, Precision, Recall, F1-Score, Accuracy**; obtained accuracy as **0.99, Pseudo R² 0.9034 (Logit) and 0.8942 (Probit)**.

SELF PROJECT

Forecasting SBI Stock Price with Time Series Analysis

(July'20-July'20)

- Forecast SBI Stock Price with previous 5 years data, checked **Stationarity, Seasonality, Trend**, using **Dickey-Fuller test, ACF, PACF** and **Decompose** plot
- Applied Time Series models – **Exponential Smoothing, AR, MA, ARMA, ARIMA, SARIMA, Holt Winter's Exponential Smoothing, VAR**
- Predicted stock price for next **22 working days (1 month)**, based on Best tuned model **ARMA (p=2, q=1)** with **RMSFE 6.41**

COURSEWORK AND SKILLS

Academic Courses	Data mining Knowledge Discovery Probability Statistics Statistical Model - Business Analytics Advanced Statistical Model - Business Analytics* Analytics Transport Telecom Intro to computing-JAVA Operation Research Advanced Decision Models Op. Management
Skills	PYTHON (Numpy, Pandas, Sci-kit Learn, Seaborn, Matplotlib, GDAL, Statmodels, etc.) R Machine Learning Statistical Analysis Natural Language Processing Deep Learning JAVA SQL MS Office CPLEX MATLAB Octave

ACHIEVEMENTS AND CERTIFICATIONS

- **Academic Excellence Awardee** for academic year 2018-19.
- Awarded **IIT Kanpur best Incoming Sports Person 2019**.
- Secured **Silver Medal for Cricket at 53rd Inter-IIT Sports Meet** held at IIT Guwahati and won **Gold Medal** at IIT Delhi Sports Festival **SPORTECH 2019**.
- Certification: **Statistics with Python Specialization | Machine Learning by Andrew Ng | Deep Learning Specialization | SQL for Data Science**

POSITION OF RESPONSIBILITY

- **Core Team Member** of **Alumni Relations** at IME Department IIT Kanpur.
- **Teaching Assistant** - Data Mining and Knowledge Discovery.