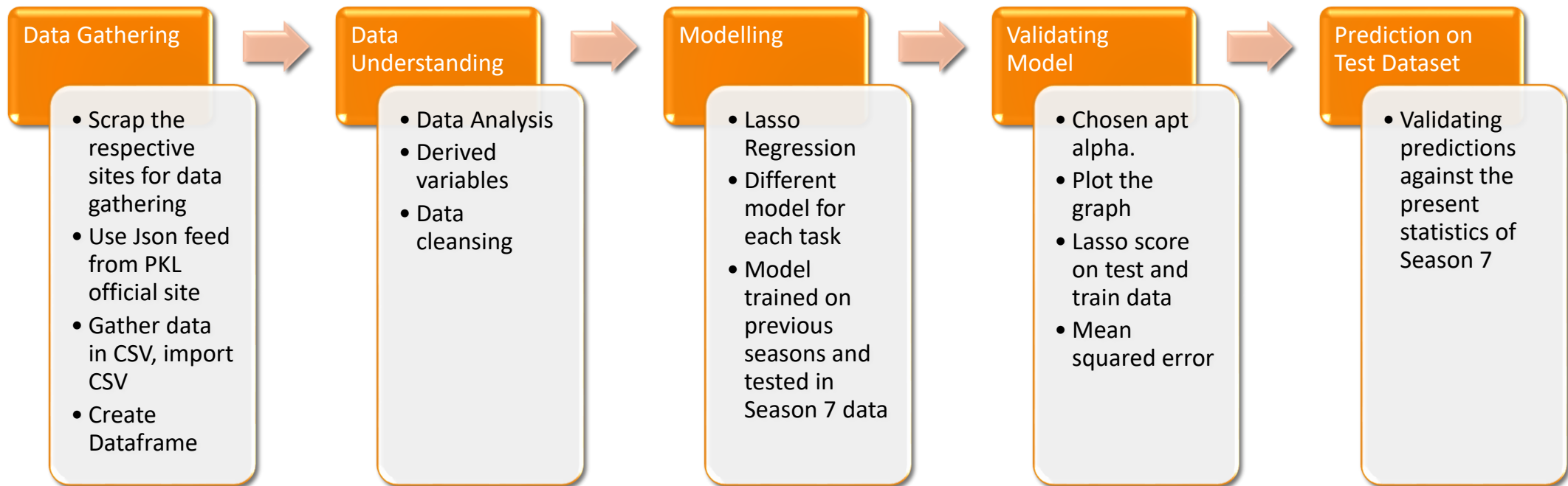# Objective

❖ To gather the data about Pro Kabaddi League from different sources by the means of copying and / or scrapping and predict the following:

  ❖ Winner of Pro Kabaddi League, Season 7

  ❖ Top team in the points table

  ❖ Top team with highest successful raids

  ❖ Top team with highest successful tackles

  ❖ Top team with highest super performance total

  ❖ Player with highest successful raid percentage

  ❖ Player with highest successful tackle percentage

Approach

Data Gathering
- Scrap the respective sites for data gathering
- Use Json feed from PKL official site
- Gather data in CSV, import CSV
- Create Dataframe

Data Understanding
- Data Analysis
- Derived variables
- Data cleansing

Modelling
- Lasso Regression
- Different model for each task
- Model trained on previous seasons and tested in Season 7 data

Validating Model
- Chosen apt alpha.
- Plot the graph
- Lasso score on test and train data
- Mean squared error

Prediction on Test Dataset
- Validating predictions against the present statistics of Season 7

Data cleansing

❖ Variables are derived wherever required from the current variables.\

❖ Dropped some of the unwanted variables which would not be influencing the model like:
  - ❖ Venue
  - ❖ Month
  - ❖ Year etc.

Python & Data files for each model

Task	Python File name	Data File
Data Extraction	Team Data Extraction.ipynb	• Used JSON feed from official website to extract the data • team_data.xlsx for information on teams • team_stats.xlsx for different statistics
Predict the top team in the points table after the completion of league matches	Task 2 - Predict_Top_Team.ipynb	• team_data_model.xlsx
Predict the team with the highest points for successful raids	Task 3 - Highest_Sucessful_Raids.ipynb	• team_detail.xlsx
Predict the team with the highest points for successful tackles	Task 4 - Highest_Sucessful_Tackles.ipynb	• team_data_model.xlsx
Predict the team with the highest super-performance total	Task 5 - Super_Performance.ipynb	• final_team_master_data.xlsx • team_stats.xlsx
Predict the player with the highest SUCCESSFUL RAID percentage	Task 6 - Highest-Sucessful Raids % - Player.ipynb	• player_details.xlsx • match_cleaned.xlsx

Python & Data files for each model

Task	Python File name	Data File
Predict the player with the highest SUCCESSFUL TACKLE percentage	Task 7 Highest-Sucessful Tackles % - Player.ipynb	• player_cleaned_data.xlsx
Predict the winner of the tournament	Task 1 - Winner_of_the_tournament.ipynb	• team_detail.xlsx

Algorithm used

❖ After much deliberation and trying different algorithms, we chose Lasso regression for our modelling, for the following reasons:

  ❖ It takes care of regularization of the data without compromising on the confidence of the model output

  ❖ It act as hyper parameter by making the least or not required coefficients to zero

  ❖ It provides a stable and robust model as compared to other algorithms

  ❖ It is quick to implement as compared to other algorithms

❖ Our r2 score is nearly 90s for both Train and Test data

Result evaluation

❖ We used the previous seasons' data (Season 1 to Season 6) to train the model and Season 7 data to test our model.

❖ The data has been fed into the model on daily basis to see if there is any change from the previous statistics.

❖ The prediction provided by the model is then compared against the statistics available on Pro Kabaddi League official website

❖ The predictions, by far, matched the present statistics of the PKL.

Predictions – Top Team

❖ Basis our model in "Task 2 - Predict_Top_Team.ipynb", Dabang Delhi seems to be the top team in the points table in league stage.

❖ We have taken successful & unsuccessful raids & tackles along with winning team and toss winner into consideration

Team
Dabang Delhi K.C.
Bengaluru Bulls
Bengal Warriors
Gujarat Fortunegiants
Puneri Paltan

Predictions – Highest Successful Raid

❖ Basis our model in "Task 3 - Highest_Sucessful_Raids.ipynb", Dabang Delhi seems to have highest successful raids.

Team
Dabang Delhi K.C.
Bengaluru Bulls
Bengal Warriors
Haryana Steelers
Puneri Paltan

Predictions – Highest Successful Tackles

❖ Basis our model in "Task 4 - Highest_Sucessful_Tackles.ipynb", Puneri Palthan seems to have highest successful tackles.

Team
Puneri Paltan
Gujarat Fortunegiants
Jaipur Pink Panthers
U.P. Yoddha
Bengaluru Bulls

Predictions – Super Performance

❖ Basis our model in "Task 5 - Super_Performance.ipynb", Bengal Warriors seems to be performing much better than others

	Team
0	bengal-warriors
2	dabang-delhi-kc
6	patna-pirates
4	haryana-steelers
10	u-mumba

Predictions – Highest Successful Raid %age

❖ Basis our model in "Task 6 - Highest-Sucessful Raids % - Player.ipynb", Pawan Sehrawat will be the most accomplishing Raider.

Playername
Pawan Sehrawat
Naveen Kumar
Pardeep Narwal
Sachin
Maninder Singh

Predictions – Highest Successful Tackle %age

❖ Basis our model in "Task 7 - Highest-Sucessful Tackles % - Player.ipynb", Faizal will out do other defenders.

Playername
Fazel Atrachali
Vishal Bhardwaj
Sandeep Dhull
Nitesh Kumar
Baldev Singh

Predictions – Tournament Winner

❖ Basis our model in "Task 1 - Winner_of_the_tournament.ipynb", Bengaluru Bulls have high chances of winning the tournament.

Thank You!