COMPREHENSIVE DATA ANALYSIS AND PREDICTION ON INDIAN PREMIER LEAGUE USING MACHINE LEARNING TECHNIQUES

[1]Dr.V.Gayatri, [2]V.Bharathi
[1,2]Associate Professor, [1,2]Department of Computer Science & Engineering, Geethanjali Institute of Science and Technology, Gangavaram, Andhra Pradesh, India

ABSTRACT
The most popular sport in India is cricket, especially the Indian Premier League (IPL), which attracts millions of fans nationwide with its T20, ODI, and Testmatch formats.Players from both regional and international teams participate in the Indian Premier League (IPL), a national cricket tournament that attracts a lot of attention from media outlets including radio, television, and live streaming. Predicting the results of IPL matches is important for online traders and sponsors, so it requires solid methods.
We used the machine learning algorithms Support Vector Machines (SVM), Random Forest Classifier (RFC), Logistic Regression, and K-Nearest Neighbor to propose a predictive model in our research. To accurately predict match results, these algorithms analyzed a variety of factors, such as team composition, player statistics (batting and bowling averages), and historical performance. Notably, the Random Forest algorithm emerged as the leader with an astounding 88.10% accuracy rate, according to our testing.
This study highlights the effectiveness of using cutting-edge computational tools to interpret complex patterns in cricket data, assisting stakeholders in making well-informed choices about the results of matches, investments, and sponsorship opportunities within the ever-changing IPL landscape.

Introduction
Cricket is an outdoor game which is played by bat and bowl which includes 2 teams of 11players each. Cricket is a teamwork game and is played mostly in three formats and occupies the 2spots in the list of the most popular sport around the World. Like in any sport, there are many factors that plays an important role in deciding the winner of the match. Selection of a team is based on the player performance and other considerations like pitch factor, team size, venue etc. There are many variables and constraints which makes The Analysis of Cricket Match Difficult.
Cricket is played in three different formats: Tests, Twenty-Twenty (T20), and One-Day International (ODI).Cricket is a national as well as an international sport.Since a single ball in cricket can decide the outcome of an entire match, every ball counts in this game.
The Indian Premier League (IPL) is a domestic cricket competition in which players are selected from India's regional teams, national squad, and international team. The entire IPL is governed by the Board of Control for Cricket in India (BCCI), and it is based on the 20-20 format and owned by celebrities, businessmen, and other individuals. Royal Challengers Bangalore (RCB), Rajasthan Royals (RR), Chennai Super Kings (CSK), Mumbai Indians (MI), Kolkata Knight Riders (KKR), Delhi Capitals (DC), Punjab Kings (PK), and SunRisers Hyderabad (SRH) are the eight teams participating in the Indian Premier League for the current season (2021).
With this, we hope to give stakeholders like club management, coaches, analysts, sponsors, fantasy cricket players, and fans accurate forecasts and insightful information. During IPL matches and tournaments, the system will have an intuitive user interface that makes it simple to interact and get real-time predictions and analysis. Using methods for machine learning such as logistic regression, random forest, SVM, etc. Because machine learning can detect fraudulent transactions, it is making our online transactions safe and secure.When we conduct an online transaction, there are a number of ways that fraud could occur, including the use of fictitious identities and accounts, as well as money theft during the transaction. In order to prevent this,FeedForwardNeuralNetworkassistance fromdetermining if the transaction is fraudulent or legitimate. The outcome of every legitimate transaction is transformed into a set of hash values, which are then used as the input for the following round. It identifies fraud and enhances our online transactions by looking for a certain pattern that changes for every legitimate transaction.

Objectives

a. Predict The First Inning Score Based on Current Situation of The Match.
b. Provide A Winner Probability For Any Condition of Second Innings of the Match.
c. Detailed Exploratory Data Analysis on Dataset.

d. Team Vs Team Analysis.
e. Player Vs Player Analysis.
f. Team past Record Analysis.
g. Player Career Analysis as Bats man and Bowler

Literature Review

Haseeb Ahmad, Ali Daud, Licheng Wang, Haibo Hong Prediction of Rising Stars in the Game of Cricket Onlinesocialdatabasesarerichsourcestoretrieveappropriateinformationthatissubsequently analyzed for forthcoming trends prediction. In this paper, we identify rising stars in cricket domain by employing machine learning techniques. More precisely, we predict rising stars from batting as well as from bowling realms. For this intent, the concepts of co-players, team, and opposite teams are incorporated and distinct features along with their mathematical formulations are presented. Generative and discriminative machine learning techniques are used for classification, and two models are assessed for each category.The suggested approach is validated experimentally as a proof of applicability, and the effects of individual features are analyzed.Additionally, model- and category-specific assessment is carried out. We show high accuracy rising star prediction that is both robust and statistically significant by using cross validation.Ranking lists for rising cricketers are finally compared with the international cricket council rankings using weighted average, performance evolution, and rising star ratings.

Haryong Song, Vladimir Shin, Moongu Jeon

This paper addresses an approach to estimating the location of a mobile node based on the range measurements of Cricket sensor network (CSN), where the coordinates of the mobile node are calculated via the method of tri lateration. There are, in general, two kinds of obstacles to be tackled and overcome in CSN: One is noisy distance measurements, and the other is the low data rates of Cricket sensors. To overcome these problems, we propose a fusion prediction-based interacting multiple model (FPB-IMM)algorithm.TheFPB-IMMalgorithmutilizesmultiplepositionmeasurements produced by trilateration and a self-tuning algorithm; it takes advantage of these multiple measurements to minimize the effect of noisy measurements and the low data rates by modifying a cycle of IMM with fusion prediction. The experimental results demonstrate that the proposed algorithm outperforms existing algorithms such as the Kalman filter and the conventional IMM.

Sarbani Roy, paramita Deyand Debajyoti Kunda

Social Network Analysis of Cricket Community Using a Composite Distributed Framework: From Implementation View point

This paper proposes an alternate ranking system based on social network metrics and theirevaluationinacompositedistributedframework.Themostimportantaspectofsocialnetworkdomainis voluminous data. In order to know key trends, predictive analysis of huge sets of raw data can beeffective. Again these analytics are generally very compute intensive. The most efficient solution ofsuchcompute-intensive analysis is to map such problems in a dispersed domain. This paper makes two significant contributions. First, a social network study of the cricket community was conducted to rank players and nations according to graph centrality metrics.Secondly, the study suggests a fully distributed system that provides infrastructure support for both graph processing and massive data analysis.An open-source framework called Hadoop is used in cloud environments to store and handle massive amounts of data.A common programming approach for handling data in a distributed fashion is called MapReduce.MapReduce alone, however, is inefficient for processing graphs.Giraphi is a substitute programming approach for Hadoop's graph processing.This paper illustrates the usefulness of the suggested framework in the process of analyzing large amounts of data and captures the significance of alternate ranking in this sport through a practical case study of social network analysis of the cricket community.

Priyanka S,Vysali K, KB Priyalyer

Score Prediction of Indian Premier League-IPL 2020using Data Mining Algorithms

One of the well-known outdoor activities with a wealth of real-world statistics data is cricket. The increasing popularity of IPL games necessitates a review of potential factors that influence match outcomes.This study examines several years' worth of IPL data, including player, match venue, team, and ball-to-ball statistics. The data is evaluated to derive a number of findings that can assist players perform better. It is centered on evaluating the results of Indian Premier League (IPL) games by utilizing current data mining algorithms to balance as well as an unbalanced dataset.This model is incredibly well-liked in predictive modeling.The first innings score in Twenty-Twenty (T20) cricket matches is currently forecasted based on the current run-rate, which is the number of runs scored per number of overs bowled. It

takes into account variables including the total number of wickets lost, the location of the game, and the toss. The Random Forest method is then used to predict the score in each innings and, ultimately, the outcome of the match. This study uses survey data to predict the 2020 Indian Premier League and data mining algorithms to analyze the results.

Prince Kansal, Pankaj Kumar, Himanshu Arya, Aditya Methaila

Player valuation in Indian premier league auction using datamining technique

A brand-new Twenty20 league, the Indian Premier League concluded its first season in 2008.In the sports world, player auctions are not a recent phenomenon.However, the Indian Premier League (IPL) was the first to implement player auctions in the game of cricket. Before this, there was no set procedure for assessing a player's performance and establishing its starting price. In this study, we develop many prediction models based on individual players' historical performance to forecast a player's selection in the Indian Premier competition, a cricket competition.Through the use of One-Day International (ODI) and T-20 variables related to both bowling and batting, we have identified several interpretable variables with explanatory power over auction values.The models that are created can assist decision-makers in determining player pay during the auction.

ShilpiAgrawal,SurajPalSingh,JayashKumarSharma

PredictingResultsofIndian PremierLeagueT-20MatchesusingMachineLearning

People in the fallage group find cricket to be the most engaging and thrilling game to watch and play. It's regarded as the most intriguing and unpredictable game. Many turn it into a multibillion dollar industry by speculating in the hopes of making money. The gaming market will always be on Hikea since spotfixing is a major worry.This paper examines the challenge of forecasting the outcome of a forthcoming Indian Premier League match based on the individual skills of each player, teamwork, and the team's improving style that is followed after every match.In this research, we offer a strategy that uses machine learning methods to predict the winning team from historical data. Three machine learning techniques, Support Vector Machine, CTree, and Naïve Bayes, were utilized, yielding accuracy rates of 95.66%, 97.58%, and 98.99%, in that order.

Harshit Barot, Arya Kothari, Pramod Bide, Bhavya Ahir, Romit KankariaAnalysisand Predictionof IndianPremierLeague

Cricketisthemostpopularsportafterfootball.IndianPremierLeagueortheIPListhemostpopular T20 domestic league in the world. Cricket involves lots of data and statistics. In the game ofcricket, several parameters can be used to predict the outcome of the game. The factors affecting acricket match can be combined with Machine Learning to predict the outcome of a match. Thisresearch has focused on analysing the features of the cricket matches in IPL. Moving further towardsthe analysis of the Indian Premier League, this paper has rated the Batsmen and Bowlers in a uniqueway based on their performance. A few crucial factors like team form and team strength in predictingthe match outcome apart from the conventional features like the toss, venue of the games etc., havebeen added. Further, a novel analysis of Batting and Bowling has been proposed based on BattingIndex and Bowling Index. Machine Learning algorithms like SVM, Logistic Regression, RandomTree, Random Forest and Naive Bayes have been applied, for match predictions. Lastly, the results,based on which algorithm gives the best accuracy, have been plotted. Decision Tree and LogisticRegressionalgorithms havegiven an accuracyover 87%and 95%respectively.

AmalKaluarachchi,S.VardeAparna

CricAI:A classification-basedtooltopredicttheoutcomein ODIcricket

Victory is the ultimate goal in any sport. In this work we address the winning factors in thesport of One Day International (ODI) cricket. Winning an ODI cricket match depends on variousfactors related to scoring as well as physical strength of the two teams. Some of the factors have beendescribed in the literature but there is scope for further research on analyzing them, especially withreference to predicting victory. Interesting factors include home game advantage, day/night effect,winning the toss and batting first. In this article, we have used artificial intelligence techniques, morespecifically Bayesian classifiers in machine learning, to predict how these factors affect the outcomeof an ODI cricket match. Based on the emerged results, we have developed a software tool called Cric AI.

Kalpdrum Passiand Nirav Kumar Pandey

Using machine learning to predict players' performance in one-day international cricket matches This study provides an approach that uses machine learning algorithms to forecast a cricket player's performance in an upcoming match. The statistical data of the Bangladesh national cricket team's players, gathered from reliable sports websites, is included in the proposed model along with feature selection algorithms like recursive feature elimination and univariate selection, as well as machine learning algorithms like linear regression and support vector machines with linear and polynomial kernels. The gathered statistical data is converted into a numerical value in order to implement those in the algorithms in order to implement the suggested model.Additionally, the previously described feature

selection techniques are used to extract the properties that are more closely associated with the final feature. The machine learning algorithms are also used to forecast runs achieved by a batter and runs that a bowler is expected to consider in the next game.The experimental configuration shows that the model provides up to91.5% precision for battersMahmudullah, a bowler, has a forecast accuracy of up to 75.3% according to Tamim, while other players' predictions are similarly accurate.As a result, this will aid in projecting players' future performance, resulting in improved squad selection for upcoming cricket matches.

Nigel Rodrigues, Nelson Sequeira, Stephen Rodrigues, Varsha ShrivastavaCricketSquad AnalysisUsing Multiple Random Forest Regression

To have a well-balanced squad in the game of cricket, it is imperative to analyse a player's performance. Since stadium circumstances vary, different tours require different lineups of players.As a result, these instructors must take into account a number of player characteristics in addition to a few other factors like the player's experience, performance under specific circumstances, and many more. This kind of data can be found in the player's career stats. The ideas of Multiple Random Forest Regression are covered in this paper in order to predict the value of the bowlers' and batsmen's traits in the current match, which will aid in choosing the players for the trip. The model will be applied to the game's ODI format.

Proposed Model

We developed predictive models to augment the current system and forecast individual player performance in future matches. The present work focuses on forecasts at the team level.The accuracy of predicting match outcomes is increased by developing algorithms to estimate how a batsman or bowler would fare based on their previous performances against particular opponents on various venues.

SYSTEM ARCHITECHTURE

Machine learning algorithms are completely dependent on data because this is the essential viewpoint that enables model training.However, a machine will be useless if it is unable to interpret the data before feeding it to machine learning algorithms.In other words, in order for a machine to solve a problem, we normally need to take care of the proper data, such as data that is in the right scale, group, and has significant qualities. Thus, the most crucial stage in the machine learning process is data preparation. Data preparation is the technique that improves our dataset's suitability for use in the machine learning process.

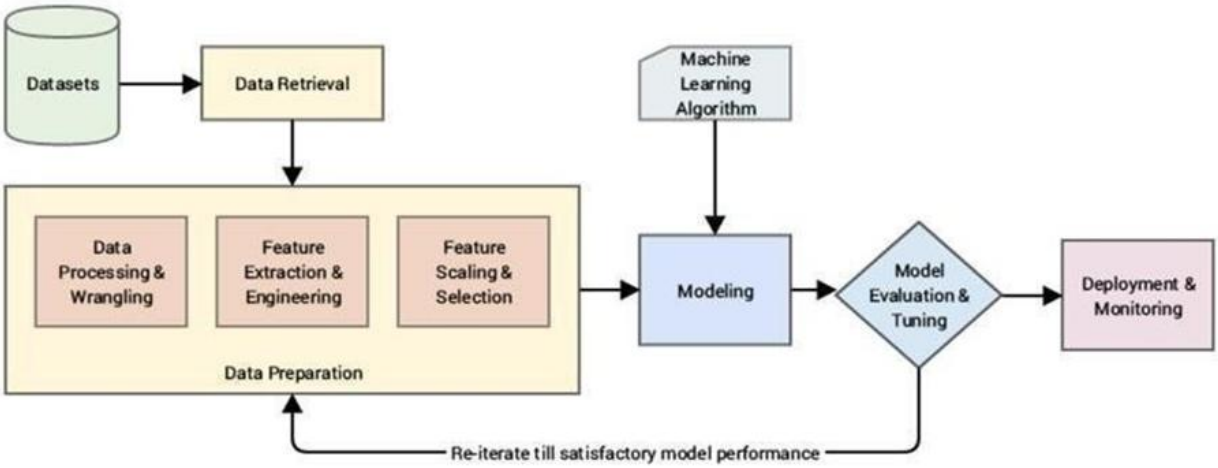

Fig.1.Work flow of the system

Figure 1 mentioned above describes the system's workflow. The following is a discussion of the system's workflow:

A dataset can be viewed as a collection of data objects, which are often also called as records, points, vectors, patterns, events, cases, samples, observations or entities.

Data objects are described by a number of features, that capture the basic characteristics of an object, such as the time at which an event occurred, etc...Features are often called as variables, characteristics, fields, attributes or dimensions.

A feature is a distinct, quantifiable attribute or aspect of an observed occurrence.Features can be numerical (interval, ratio), categorical (nominal, ordinal), or both.

DATASET

Data is essential to machine learning, and training with datasets is made possible by machine learning. The model is trained using a dataset so that it can operate automatically and perform a variety of actions.The testing dataset is what we use to verify the accuracy of our model after machine learning algorithms have been trained on the training dataset. Analyzing the dataset's numerous features will yield insightful conclusions.These elements usually include team performance data like win-loss ratio, home advantage, and historical head-to-head records, in addition to player statistics like fielding efficiency, MatchID, city, date, Toss_Winner, Toss_Decision, and winner.Advanced machine learning algorithms, like decision trees, random forests, and neural networks, are utilized to analyze these features and create models that can accurately predict player performances or match outcomes.Nevertheless, the dynamic and unpredictable nature of cricket matches also plays a role in the prediction's performance, along with the caliber of the data and the algorithms used.

	id	Season	city	date	team1	team2	toss_winner	toss_decision	result	dl_applied	winner
0	1	IPL-2017	Hyderabad	05-04-2017	Sunrisers Hyderabad	Royal Challengers Bangalore	Royal Challengers Bangalore	field	normal	0	Sunrisers Hyderabad
1	2	IPL-2017	Pune	06-04-2017	Mumbai Indians	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Rising Pune Supergiant
2	3	IPL-2017	Rajkot	07-04-2017	Gujarat Lions	Kolkata Knight Riders	Kolkata Knight Riders	field	normal	0	Kolkata Knight Riders
3	4	IPL-2017	Indore	08-04-2017	Rising Pune Supergiant	Kings XI Punjab	Kings XI Punjab	field	normal	0	Kings XI Punjab
4	5	IPL-2017	Bangalore	08-04-2017	Royal Challengers Bangalore	Delhi Daredevils	Royal Challengers Bangalore	bat	normal	0	Royal Challengers Bangalore
5	6	IPL-2017	Hyderabad	09-04-2017	Gujarat Lions	Sunrisers Hyderabad	Sunrisers Hyderabad	field	normal	0	Sunrisers Hyderabad
6	7	IPL-2017	Mumbai	09-04-2017	Kolkata Knight Riders	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians
7	8	IPL-2017	Indore	10-04-2017	Royal Challengers Bangalore	Kings XI Punjab	Royal Challengers Bangalore	bat	normal	0	Kings XI Punjab
8	9	IPL-2017	Pune	11-04-2017	Delhi Daredevils	Rising Pune Supergiant	Rising Pune Supergiant	field	normal	0	Delhi Daredevils
9	10	IPL-2017	Mumbai	12-04-2017	Sunrisers Hyderabad	Mumbai Indians	Mumbai Indians	field	normal	0	Mumbai Indians

Fig.2. Dataset collected from Kaggle

DATA PREPROCESSING

Data preprocessing is an information mining strategy that is used to change the unrefinedinformationinanaccommodatingandappreciableformat.AfterselectingtherawdataforMLtraining,theimportanttaskisdatapreprocessing.Inbroadersense,datapreprocessingwillconvertheselecteddata into a form we can work with or can feed to ML algorithms. We always need to preprocess datasothatitcanbeaspertheexceptionofthemachinelearningalgorithm.Inotherwords,featuresofthedatacan now beeasily interpreted bythealgorithm.

Need of Data Preparation and Preprocessing:

The correct formatting of data is necessary to improve the outcomes of the machine learning model that is being used.Since the Random Forest algorithm, for instance, does not support null values, null values must be managed from the original raw dataset in order to execute the random forest algorithm.Kaggle gives you access to preprocessed data. However, the preprocessing of this data is covered below.

Steps involved in Data Preprocessing:

a. Data Quality Assessment
b. Feature Aggregation
c. Feature Sampling
d. Feature Encoding

ALGORITHMS

e. In our world today, machine learning is the newest buzzword. Machine learning is being used in many different fields, including as facial identification on social media, disease recognition in the healthcare sector, weather prediction, and customer behavior analysis. In actuality, it refers to the process of feeding the machine intelligence without the need for explicit programming, analyzing historical data to make predictions about future data.Machine learning is also being used to detect malware.The term "supervised machine learning" refers to the situation in which all of the information is labeled, or, to put it another way, the knowledge of how all of the information is classified.

Regression and classification are two forms of supervised learning in which the values in the regression and the classes in the classification are known to us. Unsupervised learning, on the other hand, does not give past data; instead, associations or clustering must be done in order to create classes.
In this study, five supervised machine learning algorithms were examined. Precision, recall, and Area under theROC curve (AUC) were used as measures. We compared each method under consideration using these metrics, and in the

end, we predicted the effective model to identify credit card frauds.

Now, the algorithms we used are described below

1. Logistic Regression:
2. K-nearest Neighbors
3. Random Forest
4. Support Vector machine
5. Linear Regression

ARCHITECTURAL DESIGN

Designing an architectural frame work for IPL Match Predictor using Machine Learning models involves several steps. Below is a high-level architectural design for such a system:

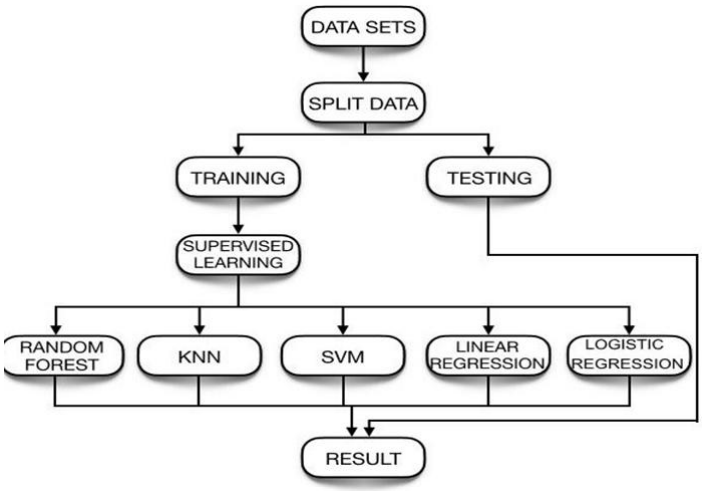

Fig.3. Architectural Design

A data-flow diagram is a way of representing a flow of data through a process or a system. The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no control flow —there are no decision rules and no loops.

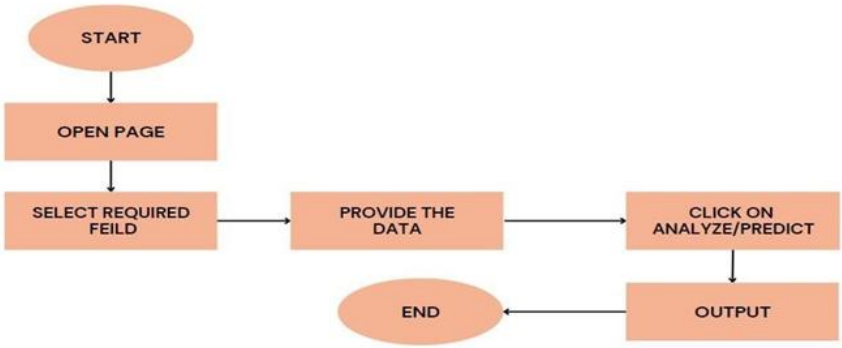

Fig.4. Data Flow Diagram

IMPLEMENTATION

SYSTEM MODULES

Data Collection

Gather datasets from sources like Kaggle.

Datasets required are matches, team statistics, player profiles, ball-by-ball data.

Data Preprocessing

Clean the data by handling missing values, formatting issues.

Feature engineer new attributes like home/a way win percentages.

Split data into train and test sets.

Model Training:

Explore different machine learning algorithms like Random Forest, SVM, Logistic Regression etc.
Train models on the training data with relevant attributes.
Tune hyper parameters to improve model accuracy.

Model Evaluation:

Evaluate models on the test set using accuracy, precision, recall etc.
Compare performance of different models.
Select the best performing model for deployment.

Model Deployment

Integrate the trained model into a prediction application/API.
Allow model to take in new match data and return predicted winner.
Re-train model periodically on latest match data.

User Interface

Create a UI to show match predictions to the end user.
Allow user to query for predictions on upcoming matches.
Visualizations to explain prediction rationale.
Here we are using stream it interface to communicate with the user.

Operationalization:

Monitor model in production, collect feedback.
Regularly re-train model on latest data.

Maintain thresholds for re-training/re-deployment

EVALUATION CRITERIA

The accuracy, precision, recall, and F1 score are the three criteria that can be used to determine the efficiency of a four model.

Accuracy

Accuracy is one metric for evaluating classification models. Accuracy is the fraction of predictions our model got right. Accuracy has the following definition:

Precision

Precision is a good evaluation metric to use when the cost of a false positive is very highand the cost of a false negative is low.

Recall (Sensitivity)

Recall is a metric that measures how often a machine learning model correctly identifies positive instances (true positives) from all the actual positive samples in the dataset. Recall is ametric that measures how often a machine learning model correctly identifies positive instances (true positives) from all the actual positive samples in the dataset. Recall can also be called sensitivity or true positive rate. The term “sensitivity” is more commonly used in medical and biological research rather than machine learning.

Confusion Matrix

When describing the performance of a classification model, also known as a "classifier," on a set of test data for which the true values are known, a confusion matrix is a softenable table.

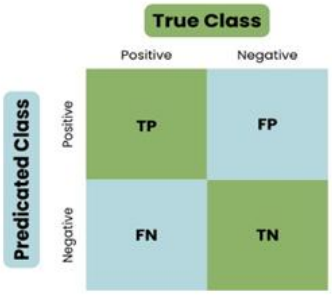

Fig.5. Confusion matrix

TESTING& VALIDATION

The goal of testing is to identify mistakes. Testing is the practice of attempting to identify every potential flaw or vulnerability in a piece of work. It offers a means of testing the functionality of individual parts, assemblies, subassemblies, and/or final products. It is the process of testing software to make sure it satisfies user expectations, meets requirements, and doesn't malfunction in an unacceptable way. Different test kinds exist. Every test type focuses on a certain testing need.

Fig.6. Home page of IPL Prediction

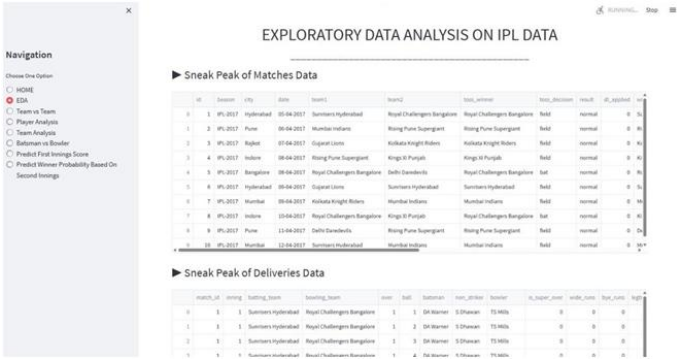

Fig.7. Exploratory Data Analysis on IPL Data

Fig.8. Team vs Team predictions

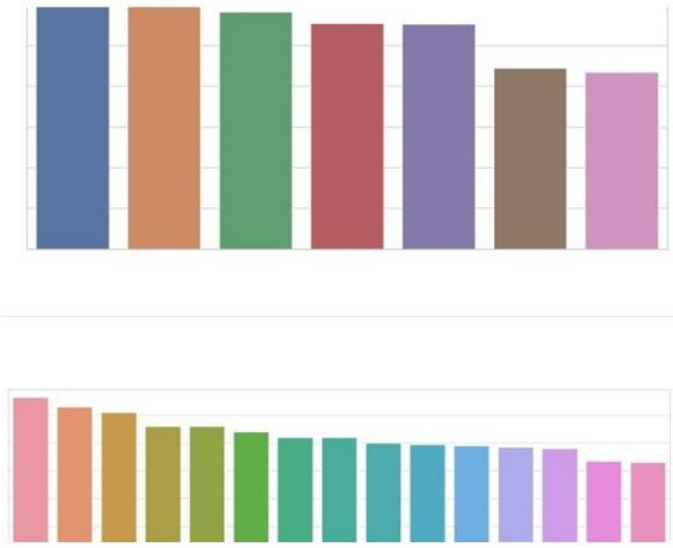

Fig.9. Player Analysis

Fig.10. Batsman vs Bowler predictions

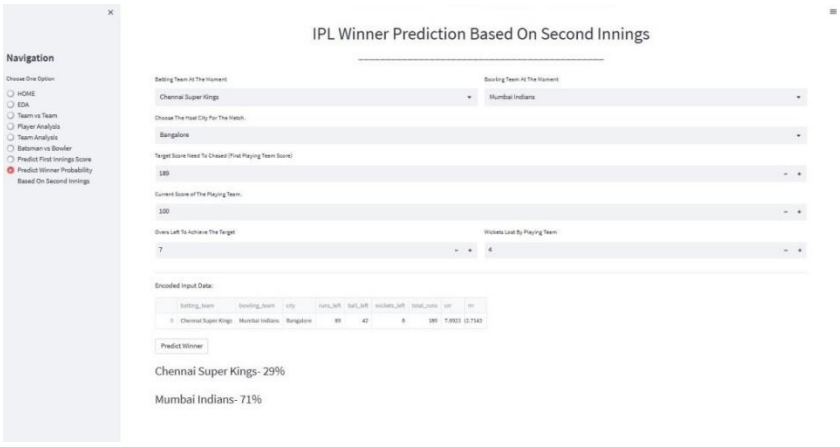

Fig.11. IPL Winner Prediction Based on Second Inning

Accuracy achieved by the algorithms

Algorithm	Accuracy
RandomForest	88.10%
K-NearestNeighbor	49.34%
LogisticRegression	51.40%
SupportVectorMachine	32.6%
LinearRegression	44.05%

CONCLUSION

The goal of this was to use machine learning techniques to forecast cricket matches in the well-known Indian Premier League event with accuracy. Because cricket is such a complex and dynamic sport, predicting the outcome of a match is a very difficult task.But thanks to the abundance of rich historical match data, machine learning models are now able to detect important variables that affect match outcomes and produce accurate forecasts.
This effort trained supervised models to predict match wins by utilizing a variety of IPL datasets that included match information, team statistics, player profiles, and ball-by-ball data.To learn more about the variables that influence match outcomes, such as team head-to-head records, a thorough exploratory analysis was conducted.To create the best prediction models, five different algorithms were assessed: Random Forest, Support Vector Machine, Logistic Regression, K-Nearest Neighbors, and Linear Regression. Random Forest was the most accurate model out of all of them, with an outstanding prediction accuracy of 88.1% on test data.The match site, player performance metrics, team compositions, and winning toss decision were found to be the most important elements in determining the outcome of the match.This demonstrates how well Random Forest modeling works to predict future match outcomes with high reliability by utilizing these critical factors from historical data.In summary, this study adds significantly to the field that combines machine learning with sports analytics.
Even though it might be very difficult to forecast the outcome of a cricket match, this study's Random Forest model demonstrates that past data can be used to generate precise predictions. The methods and knowledge gained from this research could also be applied to forecast results for other well-known athletic events. Future research could improve match outcome forecasts and facilitate fantasy cricket management by developing distinct predictive models for player performance. Overall, this work offers a strong, data-driven machine learning method for predicting cricket match winners, creating a wealth of opportunities at the nexus of AI and sports.

FUTURESCOPE

Although this study creates a useful model for forecasting the outcomes of IPL matches, there are a number of encouraging avenues for further development of this research.
Developing predictive models to anticipate individual players' performance in upcoming matches is one area of future research. The present work focuses on forecasts at the team level.Creating algorithms to estimate how a batter or bowler would perform based on their prior results against particular opponents on various fields could improve the accuracy of match prediction.
Furthermore, experiments with more sophisticated machine learning methods, such as deep neural networks, might be conducted to simulate the intricate dynamics of cricket.Predictions may be enhanced by deep learning algorithms that automatically identify relevant patterns from unprocessed player and match data. Additionally, hybrid models combining several methods could be created. Predictions can be improved by incorporating additional data sources. Information on player injuries, team tactics, pitch, and meteorological conditions may offer vital extra cues for more accurate match result forecasts. The models may also benefit from structured and unstructured data from news on team morale and social media analytics.Machine learning could be used to suggest the best team lineups, batting orders, and bowling adjustments for upcoming games in addition to match outcome forecasts.
This would increase the usefulness of AI-based cricket analytics for use by fantasy cricket sites and franchises in the real world. All things considered, a great deal of innovation at the nexus of machine learning and cricket has been made possible by this groundbreaking study.There are intriguing prospects to create even more potent AI-based forecasting

systems for this well-liked sport by creating reliable predictive models for player performance, experimenting with cutting-edge deep learning architectures, and leveraging different data sources.

References

[1] Haryong Song, Vladimir Shin and Moongu Jeon, Mobile Node Localization Using FusionPrediction-BasedInteracting Multiple Model in Cricket Sensor Network, IEEE Transactions onIndustrialElectronics, Volume: 59, Issue: 11, November2012.

[2] Priyanka S, Vysali K, K B PriyaIyer, Score Prediction of Indian Premier League- IPL 2020using Data Mining Algorithms, International Journal for Research in Applied Science & EngineeringTechnology(IJRASET),Volume 8,IssueII, PP.790-795.

[3] Harshit Barot, Arya Kothari, Pramod Bide, Bhavya Ahir, Romit Kankaria, Analysis andPredictionofIndianPremierLeague,2020InternationalConferenceforEmergingTechnology(INCET),5-7 June2020.

[4] M. B. Wright, Scheduling fixtures for New Zealand Cricket, IMA Journal of ManagementMathematics16, PP. 99−112, 2005.

[5] BalasundaramA,AshokKumarS,JayashreeD,MageshKumarS,DataMiningbasedClassification of Players in Game of Cricket, proceedings of the International Conference on SmartElectronicsandCommunication(ICOSEC2020),IEEEXplorePartNumber:CFP20V90-ART;ISBN:978-1-7281-5461-9.

[6] Do, Dinh-Thuan, Tu Anh Le, Tu N. Nguyen, Xingwang Li, and Khaled M. Rabie. "JointimpactsofimperfectCSIandimperfectSICincognitiveradio-assistedNOMA-V2Xcommunications."IEEEAccess 8(2020): 128629-128645.

[7] Rajendran, Ganesh B., Uma M. Kumarasamy, Chiara Zarro, Paramesh Chari B. DivakaraChari,andSilviaL.Ullo."Land-useandland-coverclassificationusingahumangroup-basedparticleswarm optimization algorithm with an LSTM Classifier on hybrid pre-processing remote sensingimages."RemoteSensing 12, no. 24 (2020): 4135

[8] Z. Guo, K. Yu, Y. Li, G. Srivastava, and J. C. - W. Lin, “Deep Learning-Embedded SocialInternet of Things for Ambiguity-Aware Social Recommendations”, IEEE Transactions on NetworkScienceand Engineering,doi: 10.1109/TNSE.2021.3049262.