Kanaga Lakshmi R
Information Technology
Karpagam College of Engineering
Coimbatore
rkl2000srtrk@gmail.com

RECOMMENDING OPENING MOVES FOR CHESS PLAY

ABSTRACT:

Chess is one of the most sought-after indoor games played. It's not just merely a game, it also includes problem solving, reasoning, calmness under pressure, patience, creative thinking, pattern thinking, strategy thinking and much more. There are many strategies and techniques involved in playing a chess game. Recommendation engine is used to filter the data using different algorithms and recommend the most relevant technique to the users to make them win in the play. Random forest is used to predict the outcome of the chess game. Naive Bayes is also used to predict the accuracy of the dataset. As the dataset is so large Naive Bayes classifier is used for the accuracy of the data. Additionally logistic regression, Decision tree, SVM, KNN are also used for prediction. The main objective of this project is to recommend the apt moves for the players to make them win. There are many different moves used while playing the game chess. Here, we recommend players for what kinds of chess openings tend to go together and lead them to success. Chess game dataset(Lichess) is taken from Kaggle.

KEY WORDS:

Recommendation System, Chess Moves, Logistic regression, cluster, Naive Bayes, Random Forest, Decision tree, SVM

1.INTRODUCTION:

This is the era where the world leaves the age of information and entering the age of recommendation.

The internet world has now become a surrogate way to know about people and their interest. With that information, it is apparent for the recommendation system to satisfy the users' needs. It is an information gathering or filtering system that explores to predict the rating or choice a user would make on an item. The application of recommendation systems is involved in the areas of e-commerce, e-learning, e-government and e-business services. The same mechanism is also applied to a chess game. Chess is one of the most historical games played and it is playing with the same enthusiasm even now. There are so many strategies in playing a chess game that even an expert can't master it. Before all else the opening move decides how the game will be. Opening move must be made wittily. White players have an advantage in it because he/she was the first one to start a game and his/her opening move sets the play. If the opening move was perfect the game is whites'. This may be the reason why white has greater chances of winning.

Evgeny Sveshnikov said White has to survive for a win, Black - for a draw. This was actually a true fact while playing a game. On that account it's clearly understood that the opening move is very influential. Have made a recommendation engine to recommend the opening move. Collaborative filtering is used here. It is based on the rating given by the users. Here, it sees the similarity between the users/items. Same is done with this recommendation system.

In conjunction with, used machine learning algorithms to predict the accuracy of data. List of inputs have been mapped with the output and accuracy has been predicted. Some charts have been made for easy understanding of the data.

The rest of the paper is structured as **2**. Shows the related work. **3** includes all the methodologies used here. **4** displays the output of all the methods used and **5.** Shows the conclusion and the future work.

2.PREVIOUS WORK:

Ortega[2] discussed the collaborative filtering method based on naive bayes. Explained what a recommendation system is and it's categories, content-based filtering, collaborative filtering and hybrid filtering. Hybrid algorithms combine a several filtering algorithms to get a set of items that set with the preferences of the user. In collaborative filtering we got two methods. Model based and memory based. Memory based uses a rating matrix which is used in user-based and item-based recommendation[**4, 21**]. In model based approach, a model is created from the data to make recommendations. Popular implementation of model based approach is matrix factorization and its variants. These factors remain extremely abstract for the users. This paper addressed the problem by creating a probabilistic model that the final users can interpret. The model is based on a Bayesian model that combines user-based and item-based approaches. Three different approaches based on user's space, items' space and combination of both spaces are defined.[22]

Kawai[7] presents the new hybrid collaborative filtering and content-based filtering. Describes an autonomous agent WebBot: Web Robot Agent integrates with the web and acts as a personal agent and recommends the users interest. This paper experimentally proved that hybrid collaborative filtering and content based filtering is better than collaborative filtering, content-based filtering, and combined filtering technique.

Wilkins[14] research is to prospect the extent to which knowledge can replace and support search in selecting a chess move and to lay out the issues involved. constructed a program, PARADISE (PAttern Recognition Applied to DIrecting SEarch), which finds the best move in diplomatic edged middle game positions from the games of chess masters. PARADISE avoids placing a depth limit on the search.

Nasib [19] focuses on k-means clustering behavior, ways for defeating the limitations of k-means clustering. This paper also discussed the time complexity of the k-means algorithm. The contemplated algorithm gathers the value of the number of clusters to be formed automatically. It improves the chances of finding the global optimum. This system helps in removing outliers, iteration in k-means are generated. The proposed k-means clustering is attained and implemented and the time complexity of the advised system are figured in this study.[28]

Nevada[29] discussed a decision tree is a tree whose internal nodes can be taken as tests and whose leaf nodes can be taken as categories. These tests are penetrated down through the tree to get the right output to the input pattern. Decision Tree algorithms can be applied and used in various different fields. It can be used as a replacement for statistical procedures to find data, to extract text, to find missing data in a class, to improve search engines. Many Decision tree algorithms have been devised. They have different accuracy and cost effectiveness. It is also very important for us to know which algorithm is best to use.[30]

3.METHODOLOGY:

This study discussed how a recommender system can make a player win. Recommendation system is built in such a way that it recommends the best opening move for a chess game. Collaborative filtering is used here. Also some machine learning algorithms are discussed here to find the accuracy. Here it uses logistic regression, KNN, decision tree, random forest, support vector machine and naive bayes. Predictions are made using these algorithms. Our main goal is met using this recommender engine.

3.1WORKFLOW DIAGRAM:

Initially import the required dataset ie. the chess dataset. Later data preprocessing is done. In that the data gets cleaned, null values will either get removed or will get substituted. Outliers will be removed. Data gets cleaned. Next, the machine learning algorithms will be applied and the algorithms used has been discussed below. From that, the prediction is made which algorithm suits this dataset better. This is found by the accuracy score. Henceforth, the main objective is met (ie.) building a recommendation engine. To recommend the opening move, collaborative filtering is used. Finally the prediction result and the recommended moves are exhibited as output.
This workflow is depicted in (Fig **1.**)

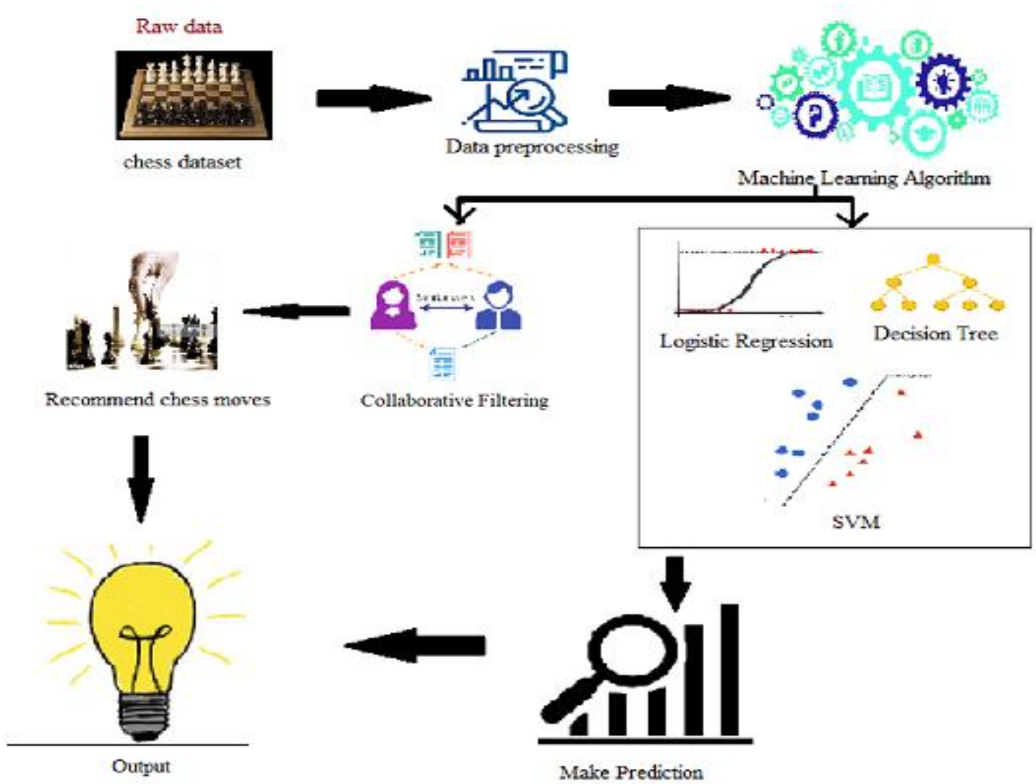


Fig 1. Workflow diagram

Logistic regression:

Logistic regression is another method obtained by machine learning from the field of statistics. It's dependent variable should be categorical. For Eg., to predict whether the email is spam(0) or not(1). Here, we have taken the winner as a dependent variable which is categorical. Logistic regression is of three types: Binary, Multinomial and ordinal. Binary has only two outcomes, Multinomial has multiple outcomes with ordering, whereas ordinal has multiple outcomes without ordering.

KNN:

K-Nearest Neighbour is a supervised machine learning algorithm used for both regression and classification. KNN assumes that similar things are close to each other. KNN arrests the idea of proximity with some mathematical concepts learned earlier. Calculating the difference between two points in a graph. However Euclidean distance given by Eq(1) is a conventional choice.

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}. \qquad \textbf{(1)}$$

Decision Tree:

Decision tree is one of the foretelling modelling approaches used in Statistics, Data Mining and Machine Learning. Decision trees are non-parametric supervised learning that uses both regression and classification. Non-parametric methods are best suitable when the data is large and we have no prior knowledge. Tree models in which we take a discrete set of values is classification tree. Classification tree is performed here.

Random Forest:

Random forest, like its name implies, consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model's prediction. Random Forest increases predictive power of the algorithm and also helps prevent overfitting. Here random forest is used to predict the accurate outcome of the game.

Support Vector Machine:

SVM is also a supervised machine learning algorithm that uses both regression and classification. Most suited for classification problems. Classification is found by the hyperplane that differentiate the classes very well. Best hyperplane should be identified which segregates the class much better. Another form of identifying a better hyperplane is it's margin. A hyperplane with high margin will be considered as a better one. Before selecting with a margin SVM selects with accuracy. Outliers will be eliminated by SVM. These are all linear hyperplanes. Non-linear hyperplanes use the kernel method. Here, linear method have been used.

Naive Bayes:

Naive Bayes is a collection of classification algorithms based on Bayes theorem. The cardinal assumption of Naive Bayes is that each feature makes an independent and equal contribution to the outcome. It is extremely fast when compared to other classification algorithms. Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred. It's given by Eq(2).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2}$$

RECOMMENDATION ENGINE:

Recommendation engine uses data analysis techniques to figure out the items that match the users taste and preferences. The Recommender system has two prominent approaches, they are content-based filtering and collaborative filtering. Content-based filtering needs a satisfactory amount of information about the item recommending. For Eg., to recommend a movie, based on content-based filtering, initially about the movie should be known thoroughly(it's actor, genre, director, year etc) or the textual content of the article. Collaborative Filtering, in different circumstances, doesn't need any other information except the historical preferences of the items. The core assumption of this approach is that those who have admitted in the past will also admit in the future. This approach is achieved through rating systems. **Explicit rating:** Rate given by the user on a particular item like 5 star rating system. It is used in many occasions. Continuation with the previous example in some cases users will be asked to rate a movie on 5 star, and the other one is **Implicit rating:** Preference made indirectly not by rating like page views, likes given to a particular item and so on. Here, a simple collaborative filtering recommendation engine is used to recommend players for what chess openings tend to win the match.

3.2ALGORITHM:

Step-by-step procedure for the proposed analysis is given.

ALGORITHM : Chess dataset.

1. Load the chess dataset

1.1. Import the required packages and clean the dataset.

2. Preprocessed the data

3. Visualized the data

4. Predicted the winner by giving victory_status and rated as input

```
features=list(zip(ratedEnc,statusEnc))
model = KNeighborsClassifier(n_neighbors=3)
model.fit(features,winnerEnc)
predicted= model.predict([[1,0]]) # 1:True, 0:draw
print(predicted)    #draw
```

Hamming Distance

$$D_H = \sum_{i=1}^{k} |x_i - y_i| \qquad \textbf{(3)}$$

5. Split the training and testing data

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)

6. Building Models

6.1. Logistic Algorithm:

Built a model
Train the model using training sets.
Predict using a testing set.
Comparing actual response values (y_test) with predicted response values (y_pred) to find the accuracy.

6.2. KNN:

Built a model
Train the model using training sets.
Predict using a testing set.
Comparing actual response values (y_test) with predicted response values (y_pred) to find the accuracy.

6.3. Decision tree:

Built a model
Train the model using training sets.
Predict using a testing set.
Comparing actual response values (y_test) with predicted response values (y_pred) to find the accuracy.

6.4. Random Forest:

Built a model
Train the model using training sets.
Predict using a testing set.
Comparing actual response values (y_test) with predicted response values (y_pred) to find the accuracy.

6.5. Support Vector Machine:

Built a model
Train the model using training sets.
Predict using a testing set.
Comparing actual response values (y_test) with predicted response values (y_pred) to find the accuracy.

6.6. Naive Bayes:

Built a model
Train the model using training sets.
Predict using a testing set.

Comparing actual response values (y_test) with predicted response values (y_pred) to find the accuracy.

6.7. Recommendation System:

Built a model

recommendations = opening_used.unstack(-1).loc[:, 'times_used'].apply(
lambda srs: srs.map(lambda v: threshold_map(v, srs.sum())),
axis='columns'

Find the cosine similarity between two users(or items) is given by Eq(4)

$$s_u^{cos}(i_m, i_b) = \frac{i_m \cdot i_b}{\|i_m\| \, \|i_b\|} = \frac{\sum x_{a,m} x_{a,b}}{\sqrt{\sum x_{a,m}^2 \sum x_{a,b}^2}} \quad \textbf{(4)}$$

Predict the model

Identify what openings a simple classifier recommends the most.

7. Make predictions from the model.

8. Display output.

4. OUTPUT:

4.1. Predict who will be the winner:

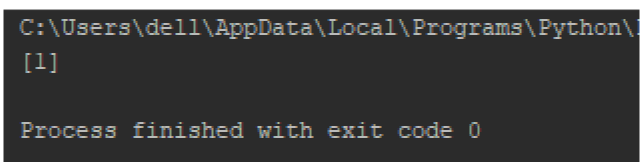
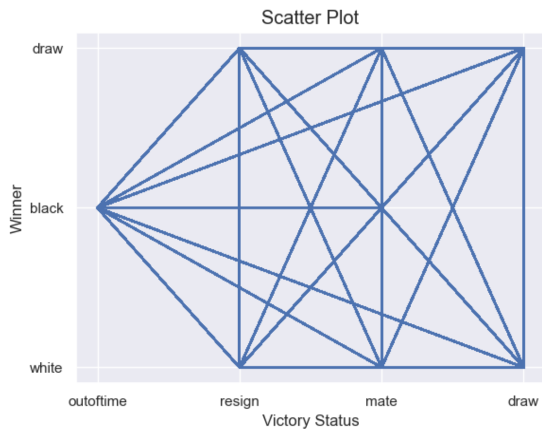


Fig 2. Winner

Algorithm discussed in **3.2(4).** In this algorithm categorical values have been encoded to numerical values. Winner is encoded as winnerEnc(white-2, draw-1, black-0). Victory_status is encoded as statusEnc(2-Outoftime,3-resign,1-mate,0-draw).Rated is encoded as ratedEnc(0-false,1-true)

In features ratedEnc and statusEnc are zipped as a list and using KNeighborsClassifier a model has been builded. Fitted features and winnerEnc in the model. So now predicting the model with rated and status it will give the winner as the predicted value. Have given rated as 1 ie. True and status as 0 ie. draw, now our result will also be draw ie.1 and we got that exactly.

4.2. Scatter Plot:



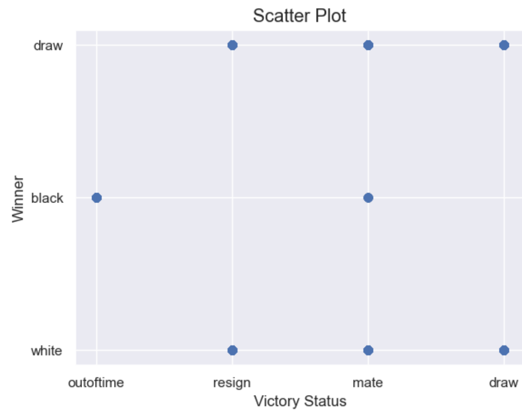


Fig 3.1. Scatter plot

Fig 3.2 Scatter plot

This is simply a scatter plot made between victory status and winner. In scatter plots we usually compare two variables looking for correlation or group. Here, the correlation between victory status and winner is obtained.

4.3. Countplot:

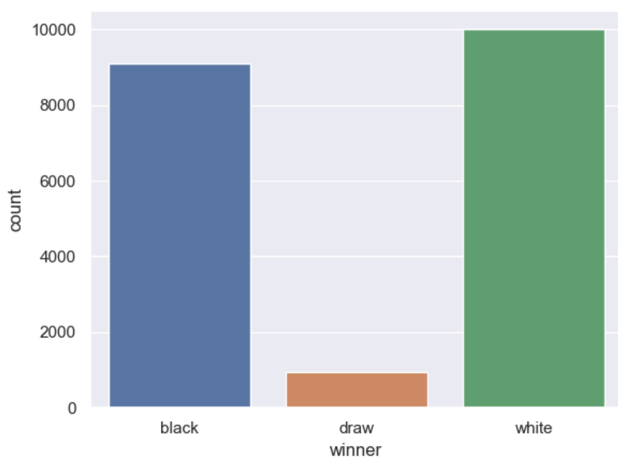


Fig 4. Countplot

Countplot displays the frequency of a given variable X using bars. Countplot works with categorical variables. Here, the x variable given is the winner. Comparatively white has more number of winnings. Draw match is very less in count.

4.4. catplot:

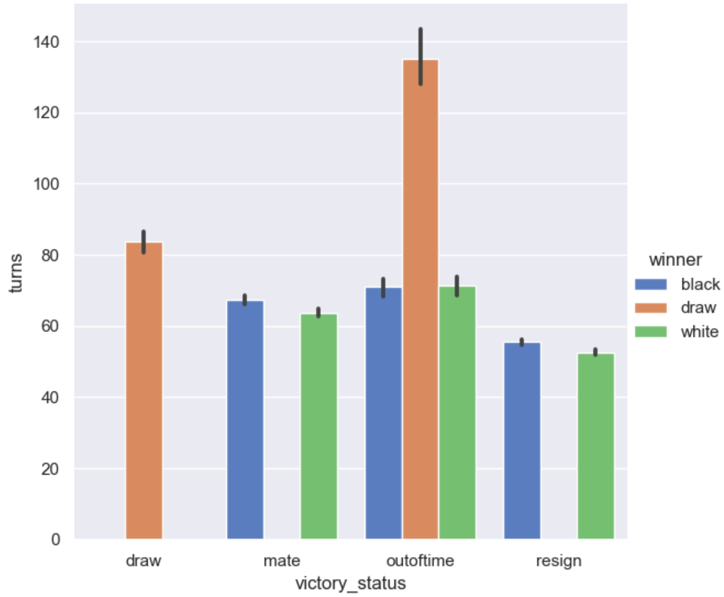


Fig 5. Catplot

The catplot function provides a new framework giving access to several types of plots that show relationships between numerical variables and one or more categorical variables. Here, categorical variable victory_status is taken as x-axis and the numeric variable turns is taken as y-axis. Filled the graph with winner(hue='winner'). In this catplot, the kind is given as bar(kind="bar"). In the draw match, the maximum number of turns have been taken.

4.5. relplot:

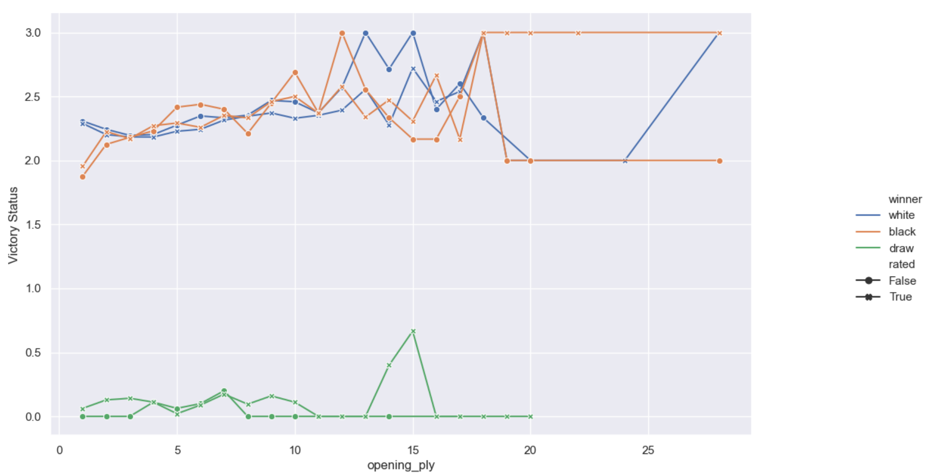


Fig 6. Relplot

In relplot the opening ply and victory status are plotted in two dimensions and in another dimension winner is added as a colour and rated as size (hue='winner',size='rated'). Here, the categorical variable victory status is changed to numerical one. (0-->draw, 1-->mate, 2-->Outoftime, 3-->resign). We will be using relplot for two numerical values.

4.6. Jointplot:

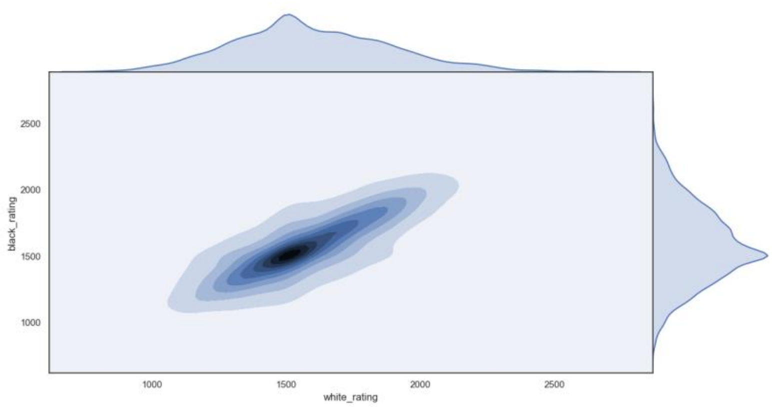


Fig 7. Jointplot

Seaborn's **jointplot** displays a relationship between 2 variables (bivariate) as well as 1D profiles (univariate) in the margins. In this plot kde(Kernel Density Estimation) is used to estimate the probability density function of a continuous random variable. It is used for non-parametric analysis. In this kde plot is used for white_rating and black_rating.

4.7. Opening Move:

4.7.1. Opening Move for Queen's Gambit Accepted

```
opening_archetype
Sicilian Defense           0.106195
French Defense             0.077015
Queen's Pawn Game          0.069840
King's Pawn Game           0.044726
Scandinavian Defense       0.043291
Queen's Gambit Declined    0.039464
Ruy Lopez                  0.038986
Italian Game               0.037312
English Opening            0.034681
Queen's Gambit Refused     0.032050
Name: times_used, dtype: float64
```

Fig 8.

Players who start the move with Queen's Gambit Accepted will also start their game with the above mentioned opening move.

4.7.2. Opening Move for Italian Game

```
opening_archetype
Sicilian Defense        0.140968
French Defense          0.069355
Queen's Pawn Game       0.063441
Ruy Lopez               0.062688
King's Pawn Game        0.052688
Philidor Defense        0.042903
Scandinavian Defense    0.035699
Scotch Game             0.032581
Four Knights Game       0.030538
Caro-Kann Defense       0.028817
Name: times_used, dtype: float64
```

Fig 9.

Players who start the move with Italian Game will also start their game with the above mentioned opening move.

From the above mentioned two opening moves there are many differences. Queen's Pawn Game and French Defense is most likely favorable for Queen's Gambit Accepted. Likewise King's Pawn game is popular with Italian Game. The Four Knights Game is one of the most popular opening moves played by experts. It is played by Italian Game players.

4.7.3. What openings does our simple classifier recommend most often?

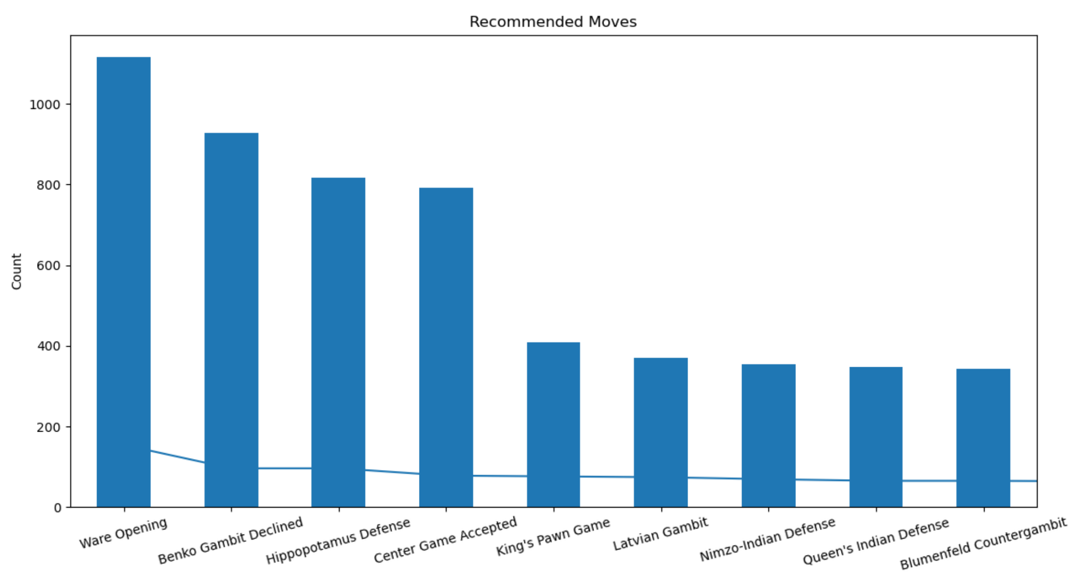


Fig 10. Opening moves

These are the opening moves that our simple recommendation engine provides. Collaborative filtering is used here as it is the simplest recommendation engine. In collaborative filtering recommendations are made by the similar users rating. Ratings can be of two types. Here, with implicit rating and with cosine similarity the similarity of items/users are found. To predict dot product is used. Finally, plotted the recommended moves that our engine provides.

4.8. Accuracy Score:

	ALGORITHM	ACCURACY
1	logistic	0.657195
2	KNN	0.689431
3	Decision Tree	0.861581
4	Random Forest	0.735294
5	SVM	0.659000
6	Naive Bayes	0.657652

Fig 11. Accuracy

The accuracy score for all the above mentioned algorithms has been displayed. Here the prediction is made on the winner of the match where the input given are victory_status, opening_move and turns. From this winner has to be predicted. With Random Forest, 74% accuracy is drawn and the highest accuracy is from Decision Tree (86%).

5.CONCLUSION AND FUTURE WORK:

Machine Learning is an application of Artificial Intelligence(AI) which makes the machines learn by itself. There are numerous algorithms in machine learning. ML algorithm is found in almost all of the things used today. For eg., the shows we see on Netflix, amazon prime, the search made in e-commerce websites, the share market predictions and so on. All these use some ML algorithms. Similarly for the chess dataset, some algorithms are used and the Recommendation System is used to recommend the opening move. This will help the players to win the play. This is actually a simple recommendation system built for the chess move. In future, further enhancements can be made to this to make it more accurate. For now this will be an eye-opener for those who will work on it later.

REFERENCE:

[1] Nguyen, T. T. S. (2019, February). Model-based book recommender systems using Naive Bayes enhanced with optimal feature selection. In *Proceedings of the 2019 8th International Conference on Software and Computer Applications* (pp. 217-222).

[2] Valdiviezo-Diaz, P., Ortega, F., Cobos, E., & Lara-Cabrera, R. (2019). A collaborative filtering approach based on Naïve Bayes classifier. *IEEE Access*, *7*, 108581-108592.

[3] Choudhary, S., Koul, S., Mishra, S., Thakur, A., & Jain, R. (2016, October). Collaborative job prediction based on Naive Bayes classifier using python platform. In *2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)* (pp. 302-306). IEEE.

[4] Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In *The adaptive web* (pp. 325-341). Springer, Berlin, Heidelberg.

[5] Felfernig, A., Jeran, M., Ninaus, G., Reinfrank, F., Reiterer, S., & Stettinger, M. (2014). Basic approaches in recommendation systems. In *Recommendation Systems in Software Engineering* (pp. 15-37). Springer, Berlin, Heidelberg.

[6] Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence* (Vol. 3, No. 22, pp. 41-46).

[7] Kawai, M., & Nogami, S. (2016). A hybrid recommender system of collaborative and content based filtering. *International Information Institute (Tokyo). Information*, *19*(6B), 2177.

[8] Borràs, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, *41*(16), 7370-7389.

[9] Burns, B. D. (2004). The effects of speed on skilled chess performance. *Psychological Science*, *15*(7), 442-447.

[10] Kister, J., Stein, P., Ulam, S., Walden, W., & Wells, M. (1957). Experiments in chess. *Journal of the ACM (JACM)*, *4*(2), 174-177.

[11] Adelson-Velskiy, G. M., Arlazarov, V. L., & Donskoy, M. V. (1975). Some methods of controlling the tree search in chess programs. *Artificial Intelligence*, *6*(4), 361-371.

[12] Baylor, G. W., & Simon, H. A. (1966, April). A chess mating combinations program. In *Proceedings of the April 26-28, 1966, Spring joint computer conference* (pp. 431-447).

[13] Simon, H., & Chase, W. (1988). Skill in chess. In *Computer chess compendium* (pp. 175-188). Springer, New York, NY.

[14] Wilkins, D. (1980). Using patterns and plans in chess. *Artificial intelligence*, *14*(2), 165-203.

[15] Barnes, D. J., & Hernandez-Castro, J. (2015). On the limits of engine analysis for cheating detection in chess. *Computers & Security*, *48*, 58-73.

[16] Janko, V., & Guid, M. (2016). A program for Progressive chess. *Theoretical Computer Science*, *644*, 76-91.

[17] Ball, T., Burckhardt, S., de Halleux, P., Musuvathi, M., & Qadeer, S. (2011). Predictable and progressive testing of multithreaded code. *IEEE software*, *28*(3), 75.

[18] Balfer, J., Bajorath, J., & Vogt, M. (2017). Compound Classification Using the scikit-learn Library. *Tutorials in Chemoinformatics*, 223.

[19].*Sudhir,Singh., and Nasib Singh,Gilll.(2013,July 7).Analysis And Study Of K-Means Clustering Algorithm.International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 7, July - 2013*

[20] Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56-58.

[21] Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In *Recommender systems handbook* (pp. 1-35). Springer, Boston, MA.

[22] Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web* (pp. 291-324). Springer, Berlin, Heidelberg.

[23] Srivastava, R., Palshikar, G. K., Chaurasia, S., & Dixit, A. (2018). What's next? a recommendation system for industrial training. *Data Science and Engineering*, *3*(3), 232-247.

[24] Harrington, P. (2012). *Machine learning in action*. Manning Publications Co..

[25] Michie, D., Spiegelhalter, D. J., & Taylor, C. C. (1994). Machine learning. *Neural and Statistical Classification*, *13*(1994), 1-298.

[26] Fagin, R., & Halpern, J. Y. (1994). Reasoning about knowledge and probability. *Journal of the ACM (JACM)*, *41*(2), 340-367.

[27] Bae, J. (2014). *Probability and statistics*. Stanford University, Stanford.

[28] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *IEEE transactions on pattern analysis and machine intelligence*, *24*(7), 881-892.

[29] Nevada, A., Ansari, A. N., Patil, S., & Sonkamble, B. A. (2011, June). Overview of use of decision tree algorithms in machine learning. In *2011 IEEE control and system graduate research colloquium* (pp. 37-42). IEEE.

[30] Zhong, Y. (2016, August). The analysis of cases based on decision tree. In *2016 7th IEEE international conference on software engineering and service science (ICSESS)* (pp. 142-147). IEEE.