

# Methodology for Analysis of Induction Motor's Design Parameters using Machine Learning

Mamidi Ramakrishna Rao

Independent Design Engineering Consultant, Visakhapatnam 530043, Andhra Pradesh, India

**Abstract** - Machine learning (ML) algorithms are being applied across multiple industries and across value chain including engineering. Typically, large size induction motors (>1 MW) are designed and manufactured to application needs. Objective of this paper is to develop a methodology for generating a technical databank for large size induction motors (>1MW) and analyze the same. This method can be further developed for empowering sales & marketing teams during customer interactions for large size motors. Making information available and accessible which can be customized to client requirements in real-time leads to efficient client interactions. ML algorithms (in Python) are used to analyze and understand induction motor parameters such as torque & current characteristics during starting and efficiency & power factor during running with active material cost. For parameter analysis, a data set is generated. The data is analyzed using 'Scikit-learn'. Data visualization is done with Pandas and Matplotlib. Classifiers are used in data analysis. The classifier accuracy is checked for accuracy with the data generated. In design data, 50 designs are shown. The number of designs can be increased for improving the performance.

**Key Words:** Machine Learning, Induction motor, performance data.

## 1. INTRODUCTION

Induction motors have various applications including driving pumps, compressors, crushers, mills, fans, etc. Selection of motor is critical for its functioning as it is required to drive the equipment. This means, the driving equipment should match the driven equipment characteristics. In addition to motor selection, it is necessary to consider equipment working environmental conditions and power system supply variations.

Typically, sales & marketing team is better equipped with information relating to small and medium rating motors. It is possible to match the equipment requirements with motor specifications by referring to manufacturer's published literature. However, for large size motors (>1 MW), which are custom made for a specific application, selection of parameters becomes critical and motor designs are application-specific. Preliminary engineering work needs to be done to set the parameters to meet customer requirements. Sales & marketing teams are unable to efficiently address and discuss client's custom requirements in real-time and requiring them the help of design &

engineering departments for preliminary calculations. This increases manpower involved, communication lags leading to delay in addressing client's requests.

In this paper, attempts are made to generate data, analyze the same and present it as a method enabling sales & marketing teams have better access to information. They can be self-sufficient in conducting preliminary calculations and meeting client's requirements before passing it on for detailed engineering work.

For calculations a 1 MW HT motor has been selected. Machine learnings (ML) techniques are used in classification and scattering the data.

**Table - 1:** Nomenclature

Symbol	Description
f	Supply frequency
I	Phase current
p	Pole pairs
P	Motor power
Scage	Single-cage rotor induction motor
DD-cage	Aluminum die-cast-double cage rotor induction motor
D-cage	Fabricated double-cage rotor induction motor
V	Supply Volts
Wound	Wound rotor with copper winding
$\omega$	Synchronous angular frequencies

## 2. APPROACH

The power ratings, in large size motor application, typically vary between 1 to 10 MW and higher. Higher power ratings, are special case requirements and custom built. To illustrate the use of ML in induction motor design selection and analysis, a 1MW, 6.6 kV high tension, 50 Hz, 4 pole motor is selected.

Depending on electrical design and rotor construction, induction motors have several types of starting torque-speed and current-speed characteristics. Commonly there are four types of rotor construction - single cage (Scage), aluminum-die- cast-double cage (DD-cage), fabricated-double-cage rotor (D-cage) and wound rotor designs (Wound). In single cage rotors, the cage material could be either copper, aluminum or alloys of copper. The rotor slots shape may be

simple rectangle or deep skin effect predominant bar. Typical torque speed and current speed characteristics in per-unit values of single, double and wound rotor motors are shown below in Figure 1 and 2. The material cost estimated is considering active materials i.e. stator and rotor magnetic lamination and copper.

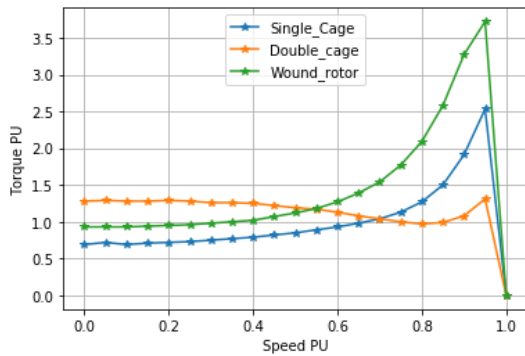


Chart - 1: Comparison of torque vs speed (typical) characteristics

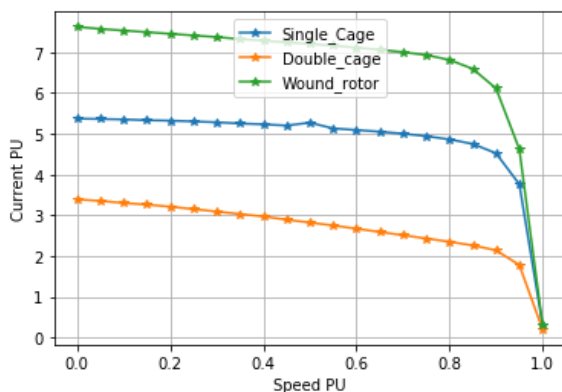


Chart - 2: Comparison of current vs speed (typical) characteristics

The ‘critical parameters’ for analysis are

- i) Starting\_Current (PU)
- ii) Torque\_Current (PU)
- iii) Maximum\_Torque (PU)
- iv) Efficiency (%)
- v) Power factor
- vi) Material cost (active) in INR

The first five parameters are calculated design outputs. For estimating the active material cost (copper and iron weights i.e. stator and rotor winding copper and lamination weights) design output weights are considered. In data visualization charts, normalized material cost is considered.

Steps followed are

Step 1 - Generate design data bank. The data bank is generated for 1 MW 6.6kV 4 pole 50 Hz by using industry

proven design and analysis software. Design details and output performance parameters are outlined in the “Annexure” section.

Step 2 - Prepare data ready for analysis.

Step 3 - Data analysis and visualization. For data analysis and visualization, python libraries, matplotlib, keras, sklearn, numpy are used.

After data generation and data preparation, available / published working python code is used. The code references are given in “Reference”. The data generated is shown in “Annexure”.

### 2.1. Machine Details:

Table 2: Machine details

Particular	Value
Rating (kW)	1000 kW
Volts (V)	6600 Volts ± 5%
Frequency (f)	50 Hz ± 0.25%
Poles (2p)	4
Synchronous speed	1500 RPM
Rotor type	
a) Cage	Copper
b) Cage	Aluminum
c) Cage	Bronze
d) Wound	Copper

In order to check the motor suitability for the drive, following the points need to be checked

- a. Matching of torque-speed characteristics of drive and driven equipment. The drive equipment is usually either centrifugal or reciprocating pumps, compressors rotary, reciprocating, centrifugal fans, blowers, crushers, mill duty and pulley drives. The starting torques characteristics are different. In case of large machines applications, the driven equipment data is provided by its manufacturer.
- b. Method of starting: For cage motors start either Direct-online (DOL) or reduced voltage starting or soft starting. This depends upon the power system strength.
- c. For wound rotor motors, to reduce the starting current in rush and improve the starting torque, controlled external resistance is inserted in rotor circuit. Performance data given for wound rotors in data bank is without external resistance.

### 3. DESIGN DATA

Published data banks are not available for motor performance. Therefore, a design data bank needs to be generated. Data bank generation is possible through two ways -

- a) From the experience of previously manufactured motors.

OR

- b) By generating new designs, using proven design software.

In the present case, designs are generated using a proven design software. The generated designs must be practical and possible to manufacture. Taking the base design data as reference, new designs are worked out. Variables considered for generating designs are magnetic core length and diameter. Constant parameters are radial air-gap, turns per phase, slot combination and insulation scheme.

### 3.1.Data Pre-processing:

In order that machine can easily parse it and the features of the data is interpreted by the algorithm, preprocessing the data is necessary. There are six steps to be followed in preprocessing of large unprocessed data. However, in our case, as we are developing data with specific purpose, the number of preprocessing steps can be reduced. Basic step in preprocessing, involves elimination of rows with missing data, removal of inconsistency in data, removal of duplicate data, etc. The second step is data-normalization and standardization. Machine learning methods expect the data attributes to have the same scale. In normalization, the attributes are in the range 0 and 1. (or -1 to +1)

### 3.2.Base design:

Designs for ‘Design bank’ are generated by taking the reference from base design. The base design has parameter values as shown in Table 3.

Rotors considered are cage and wound rotor motors. Various torques speed characteristics are obtained by changing the rotor cage material to either copper or aluminum or copper alloy. New designs are generated by varying the diameters and core lengths and rotor type (single / double cage). They are within following limits:

- a) Constant values taken from base design:
  - a. Voltage, frequency, poles
  - b. Slot combination
  - c. Radial air-gap
  - d. Turns per phase
  - e. Insulation scheme
  - f. Stamping lamination material grade

- b) Total number of designs generated – 50

**Table 3:** Base design data

Particular	Value
Stator lamination diameters (Outer/Inner)	900/580
Rotor lamination diameters (Outer / Inner)	576/290
Radial air-gap (mm)	2.0
Core length (mm)	700
Turns per phase	176
Slot combination	48/60
Rotor cage material	Copper

**Table 4:** New design parameter limits

Particular	Value
Stator laminations outer diameter (mm)	800 to 900
Core lengths (mm)	500 to 700
Rotor lamination inner diameter (mm)	240 to 290
Rotor type	Cage/ Die-cast double cage/ Fabricated double cage/ wound
Rotor cage material	Copper/ aluminum/ copper alloy

- c) Total number of designs generated – 50

### 3.3.Design Attributes:

The above identified ‘critical parameters’ are the attributes as given in approach.

## 4. ANALYSIS

### 4.1.Statistical description:

Prior to normalizing the data, it is interesting to know the statistical details such as count, mean, standard deviation (std), minimum and maximum of the data. ‘Pandas’ is used to view the statistical description.

The data is shown in Table 5. From Table 5, we can observe from “count” that there are total 50 number of designs. The performance data, “ $I_{st}$ ,  $T_{st}$ ,  $T_{max}$ , Efficiency, PF”, shows the design coverage variety.

The count plot for different types of motors is shown in chart 3

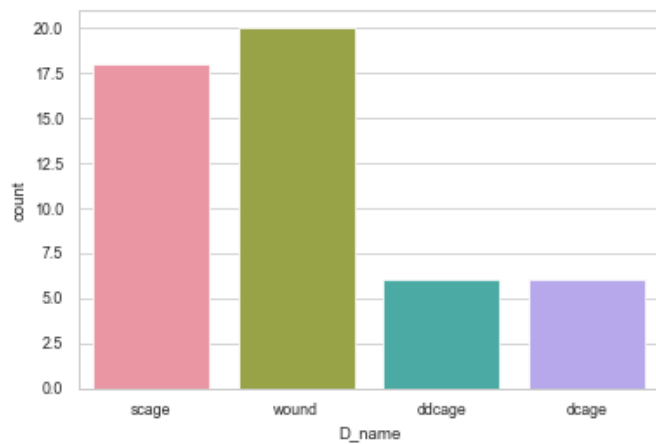


Chart 3: Count plot for motor types

Table 5: Statistical description of data in data-frame

	Ist	Tst	Tmax	Efficiency	PF	Mat_cost
Count	50.0	50.0	50.00	50.00	50.0	50.00
Mean	5.10	0.97	2.37	95.26	0.88	433138.92
Std	0.84	0.48	0.53	1.26	0.02	45172.81
Min	3.30	0.27	1.23	91.57	0.81	340960.00
0.25	4.66	0.62	1.99	95.28	0.87	400967.50
0.50	5.09	0.79	2.49	95.52	0.89	434191.00
0.75	5.60	1.45	2.70	96.10	0.90	452564.00
Max	7.83	1.88	3.75	96.54	0.92	616688.00

### 4.2.Normalization:

Normalization is a process in data preparation. Many ML methods are effective, if the data is normalized and standardized.[1] In normalization, the values are rescaled in the range of 0 and 1 (or from -1 to 1).

Chart 4 shows a scatter plot with normalized data of starting current and pull-out torque. Trend shows for higher pull-out torque requirement, larger starting current is required. Similar such scatter plots can be obtained with other variables

We can visualize the variation between three variables. Two such plots are shown below. Chart 6 shows variation of material cost with starting current and torque. Chart 7 shows the pull-out torque variation with starting current and starting torque

We have seen the visualization of the data as total data. Now let us visualize the relative performance of each class namely wound rotor and cage rotor etc.

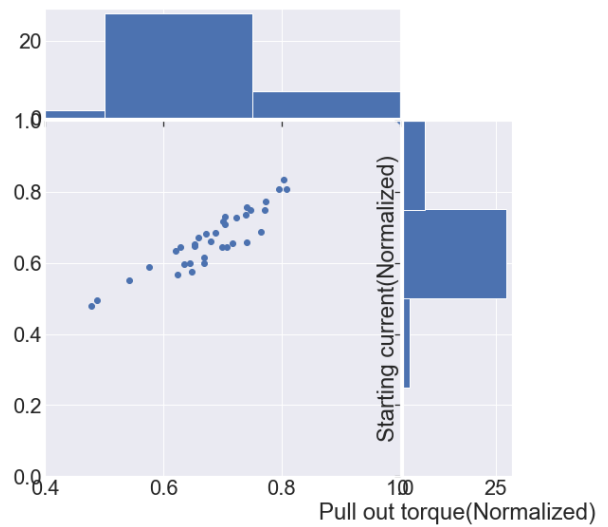


Chart 4: Scatter plot and histogram of pull-out torque and starting current (Normalized)

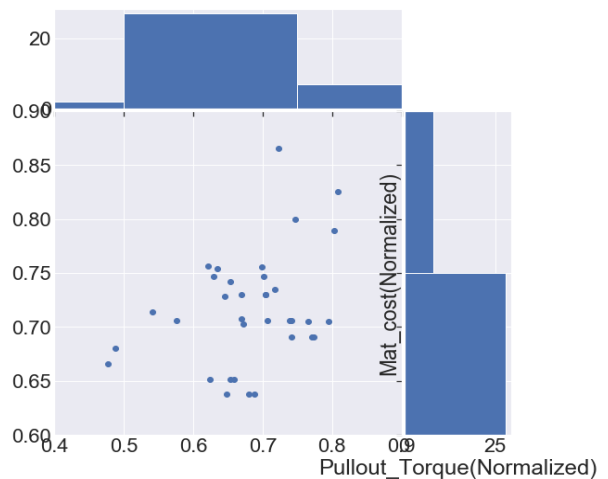


Chart 5: Scatter plot and histogram of pull-out torque and material cost (Normalized)

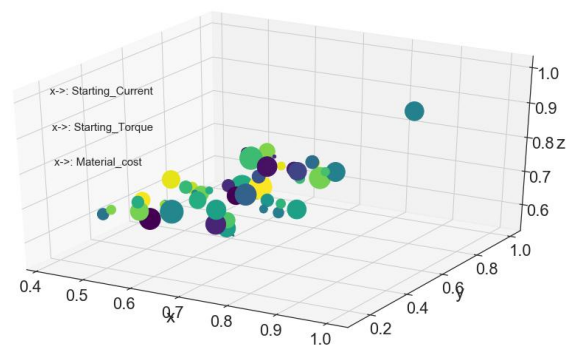
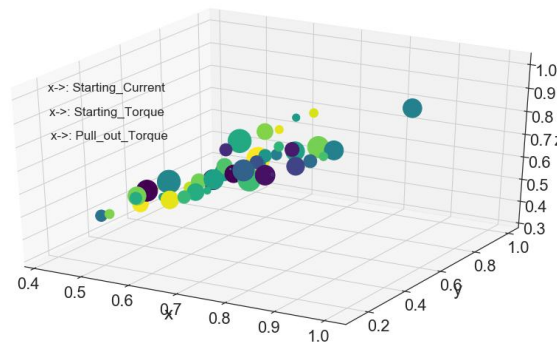
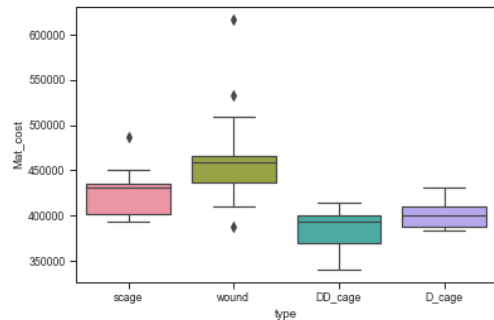


Chart 6: 3D Scatter plot of starting current and starting torque with material cost

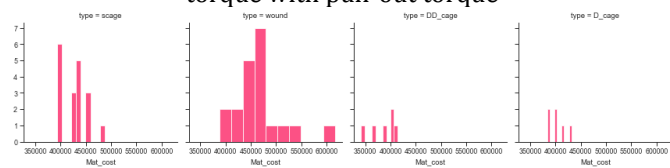
The following three charts (histogram chart, box plot and violin plot), show the material cost of cage, wound rotor, die-cast-double-cage and double- cage rotor motors).



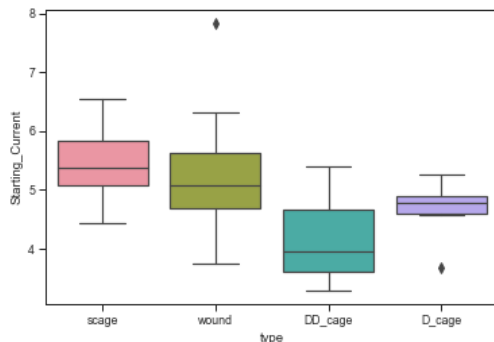
**Chart 7:** 3D Scatter plot of starting current and starting torque with pull-out torque



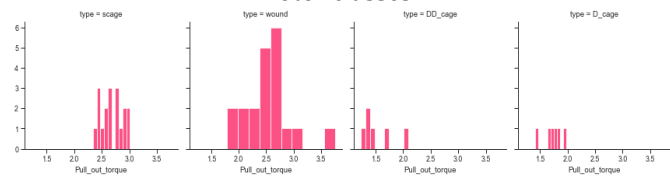
**Chart 12:** Box plot showing material cost for rotor types



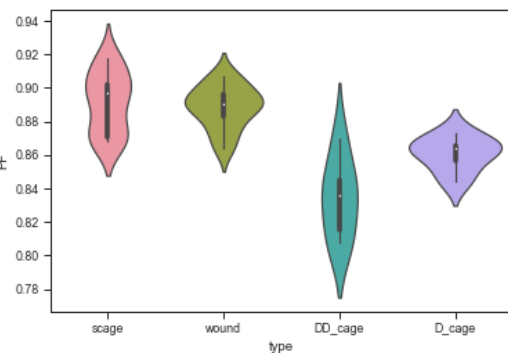
**Chart 8:** Histogram chart showing the material cost of rotor classes.



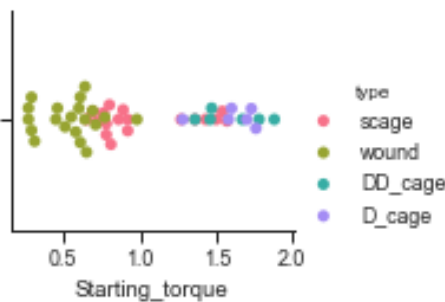
**Chart 13:** Box plot showing the Starting current for rotor types



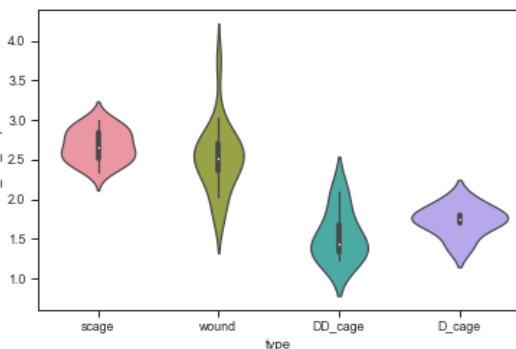
**Chart 9:** Histogram chart showing the Pullout torque of rotor classes



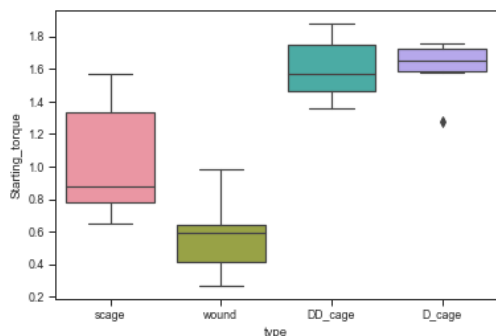
**Chart 14:** Violin plot showing the power factor for rotor types



**Chart 10:** Showing starting torque for rotor types



**Chart 15:** Violin plot showing the Pull-out torque for rotor types



**Chart 11:** Box plot showing the Starting torque for rotor types

From the above three plots, we can get the material cost variation, limits, range for both cage and wound rotor machines.

Similarly, we can draw the plots for other variables like starting current, starting torque, Maximum torque, efficiency and power factor

## 5. CLASSIFIERS

In the above data set, there are two main classes - cage rotors and wound rotors. For analysis, it is necessary to distinguish whether it belongs to “cage” or “wound” class. The classification between two classes is known as “Binary classification”. Popular algorithms [2] that can be used for binary classification include –

- a) Logistic regression classifier
- b) Decision tree classifier
- c) K-NN nearest neighbor classifier
- d) Linear Discriminant Analysis (LDA) classifier
- e) Gaussian Naive Bayes (GNB) classifier.
- f) Support Vector Machine (SVM)

### 5.1. Logistic Regression Classifier

Logistic regression is classification algorithm. It is imported from sklearn.linear\_model. It is a predictive analysis algorithm and uses ‘Sigmod function’ as cost function and hence is called ‘Logistic regression’. Taking two parameters such as ‘Pull-out torque’ and ‘material cost’, we can apply this ‘Logistic regression’ analysis tool and obtain the decision boundary

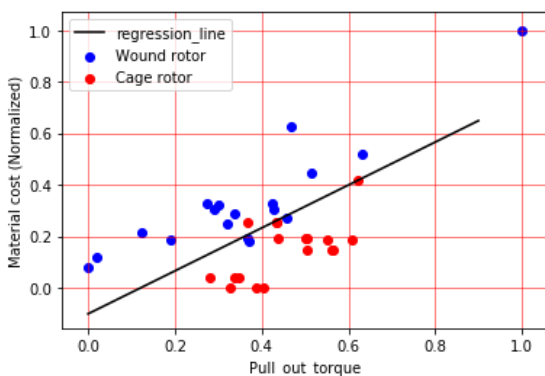


Chart 16. Scatter plot with regression analysis between ‘pull-out torque’ and ‘material cost’.

Similar such plots can be drawn for other parameters also.

The performance accuracy of the above classifiers is compared.

For working out detailed analysis and prediction, it is necessary to have training data and testing data. To avoid over fitting, it is common practice to hold out a part of

available data as “test data”. In scikit-learn, random split “training set” and “test set” is done.

The ‘Linear Discriminant Analysis’ (LDA) is imported from sklearn.discriminant\_analysis [5]. The “GNB” is imported from sklearn.naive\_bayes [6]. The ‘Decision Tree’ is imported from sklearn.tree[7]. “Support Vector Machine’ is imported from sklearn.svm.[8]

Using the train\_test\_split helper function and the algorithms given in sklearn, accuracies of all classifiers are determined.

From the data set, we have

x.shape, y.shape = (50,5) , (50)

x.train shape, y.train.shape = (37,5) , (37)

x.test.shape, y.test.shape = (13,5) , (13)

The accuracies for various classifiers for the dataset are given in Table 6.

Table 6: Accuracies of classifiers

Classifier	Set type	Value
Logistic regression	Training set	0.73
	Test set	0.85
Decision tree	Training set	1.00
	Test set	0.92
K-NN	Training set	0.89
	Test set	0.92
LDA	Training set	0.92
	Test set	0.85
GNB	Training set	0.95
	Test set	0.92
SVM	Training set	0.92
	Test set	0.92

## 6. DECISION BOUNDARY

It is necessary to plot a decision boundary [3] between classes. In the current dataset, we have two classes namely cage and wound. The following figure shows a decision boundary plot for two variables “efficiency’ and ‘material cost’. Similar such plots can be made for other variables

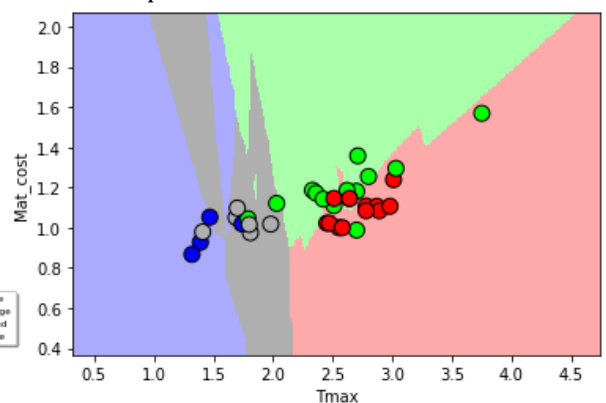


Chart 17: Decision boundary plot (efficiency and material cost)

## 6.1. Nearest Neighbors classification

'K-Nearest Neighbors' is imported from sklearn.neighbors. In KNeighbors Classifier [4], k is the number of nearest neighbors and is critical parameter which is specified by user. The accuracy of kNN depends upon the value of 'k' which in turn depends on the data. For the current dataset, accuracy is evaluated for 'k' values up to 20 and accuracy obtained for different values k is shown chart. 18. The accuracy did not improve with increased k. The best accuracy is obtained with k= 2

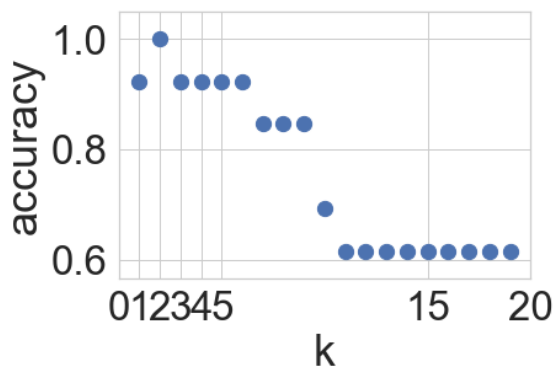


Chart 18: Calculated accuracy values for different 'k'

## 7. DISCUSSION

One example of usage of the above data is discussed here. Let us take a case where there is requirement of three motors for three applications. The power rating is nearly same. The applications are centrifugal pump, reversing re-rolling mill duty and lastly hammer mill. Now let us see, how to decide choosing the right motor and the cost, type of motors. In order to decide, first we have to understand the driven equipment and its starting and running characteristics. In the present case, the driven equipment needs are different. In case of centrifugal pump, the starting torque requirement is less of the order of 25 to 30%. After the pump picks up full speed, the load is steady.

Now, we can take the help of ML tools developed. The starting torque is only, 25% to 30% of full load torque. From "Chart 11: box plot shows the starting torque for rotor types. It is clear that, it is not necessary to choose double cage or die-cast double cage rotors. So, from starting torque point of view, both single cage or wound rotor motors are suitable. Now, to decide, whether to select wound rotor or single cage rotor, let us look other parameters starting current and cost. the starting current and cost. chart 12 and chart 13 give the needed information. Wound rotor motor cost more than single cage rotor motor. So, the decision is to select single-cage-rotor motor. After deciding the type and performance, we can refer to 2D scatter plots charts 4,5 and 3D scatter plots charts 6,7 for identifying the closet design which can satisfy the application.

In the second case - rerolling mill, this application duty calls for high pullout torque. When referring to chart 9 and 15, it is clear that the wound rotor motor is right choice. For other performance parameters, we can refer to other above charts. The third application is hammer mill application which calls for high starting torque. From charts 10 and 11 we can select either die-cast double cage rotor or fabricated double cage. chart 15 helps to decide from pull-out torque requirement. It should be remembered that die cast double cage rotors are relatively more robust than fabricated double cage rotors.

From the above data and charts, we can observe that,

- a) Data for 50 designs are generated. We can increase the number of designs for improving the overall performance.
- b) It is possible to increase the number of attributes like inertia capabilities, thermal capabilities.

## 8. CONCLUSION

The above analysis and data visualization of machine learning technique shows at a glance the data distribution and feature information. This is useful in motor application selection. The critical part is data bank generation. The designs generated must be manufacturable. Secondly, as the data is normalized, the application need not be restricted to 1 MW. It can be extended to higher power ratings as long as the trend remains same.

## 9. REFERENCES

- [1] Jason Brownlee: "Rescaling Data for Machine Learning in Python with Scikit-Learn <https://machinelearningmastery.com/rescaling-data-for-machine-learning-in-python-with-scikit-learn/>
- [2] Susan .Li "Solving A Simple classification problem with Python- 'Fruit Lovers". Link: <https://towardsdatascience.com/solving-a-simple-classification-problem-with-python-fruits-lovers-edition-d20ab6b071d2>
- [3] Scikit learn (0.23.1) " Nearest Neighbors classification" [https://scikit-learn.org/stable/auto\\_examples/neighbors/plot\\_classification.html](https://scikit-learn.org/stable/auto_examples/neighbors/plot_classification.html)
- [4] kNN classification using Scikit-learn <https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn>
- [5] LDA: sklearn.discriminant\_analysis, scikit learn [https://scikit-learn.org/stable/modules/generated/sklearn.discriminant\\_analysis.LinearDiscriminantAnalysis.html](https://scikit-learn.org/stable/modules/generated/sklearn.discriminant_analysis.LinearDiscriminantAnalysis.html)
- [6] GNB: sklearn.naive\_bayes.GaussianNB [https://scikit-learn.org/stable/modules/generated/sklearn.naive\\_bayes.GaussianNB.html](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html)

[7] Decision Tree: sklearn.tree.DecisionTreeClassifier

<https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html>

[8] Support Vector Machines (SVM) ,sklearn.svm.SVR

<https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html>

## 10. ACKNOWLEDGEMENT

The author wishes to thank Ms. Sri Rekha Mamidi for her sincere help in preparing the paper and presentation to confirm requirements. Also, the author sincerely thanks Mr. Jagdish Mamidi, for his help in code implementation.

## BIOGRAPHIES

Author has done his Master's Degree in Electrical Engineering with specialization in Rotating Electrical Machines from Indian Institute of Technology, IIT (Mumbai) in 1970. He has over 35 years of Industry experience and has worked in companies including Crompton Greaves Ltd, Kirloskar Electric co and WEG (India). He has been providing consultation to large reputed companies. His interests are in the field of design of large electrical machines, Optimization and Machine learning. He has presented technical papers in national and international conference and published in journals.

## Annexure

**Table 7: Design Data Bank**

Starting Current (pu)	Starting torque (pu)	Pull-out torque (pu)	Efficiency %	PF	Active material cost (Rs)	Rotor_type
5.37	0.69	2.58	96.38	0.868	393188	scage
5.18	0.81	2.55	95.9	0.868	393188	scage
4.51	1.36	2.43	91.77	0.87	393188	scage
5.15	1.54	2.78	92.62	0.872	426180	scage
6.06	0.78	2.9	96.54	0.87	426180	scage
5.87	0.92	2.89	96.11	0.87	426180	scage
7.83	0.98	3.75	95.44	0.893	616688	wound
4.97	0.6	2.33	95.66	0.895	466438	wound
4.31	0.51	2.03	95.28	0.892	440572	wound
3.76	0.45	1.79	94.87	0.884	410568	wound
3.88	0.46	1.83	94.85	0.889	419822	wound
4.61	0.55	2.16	95.29	0.894	435380	wound
6.32	0.8	2.98	96.53	0.884	434960	scage
5.39	1.57	2.87	92.84	0.886	434972	scage
5.7	0.69	2.71	95.37	0.898	533546	wound
6.32	0.77	3.03	95.35	0.89	508996	wound
4.82	0.29	2.51	95.52	0.864	436348	wound
5.05	0.29	2.62	95.62	0.873	466118	wound
6.54	0.89	3.01	96.34	0.918	486980	scage
5.72	0.78	2.64	96.22	0.914	450350	scage
5.54	0.92	2.64	95.72	0.914	450350	scage
4.7	1.5	2.51	91.57	0.914	450350	scage
5.14	0.31	2.69	95.61	0.864	453302	wound
4.69	0.27	2.42	95.52	0.875	449152	wound
5.26	0.65	2.47	96.28	0.897	401974	scage



5.07	0.76	2.45	95.78	0.897	401974	scage
4.43	1.27	2.34	91.66	0.897	401974	scage
5.05	1.43	2.65	92.5	0.902	435486	scage
5.75	0.86	2.77	96.1	0.901	435486	scage
5.92	0.72	2.78	96.44	0.901	435486	scage
5.86	0.71	2.8	95.3	0.891	493090	wound
5.1	0.61	2.45	95.12	0.887	457866	wound
5.05	0.58	2.36	95.31	0.907	460746	wound
5.34	0.61	2.52	95.37	0.901	433422	wound
5.61	0.64	2.63	95.45	0.905	460746	wound
4.67	0.27	2.38	95.59	0.886	465228	wound
3.56	1.47	1.32	95.05	0.809	340960	DD_cage
3.78	1.46	1.39	95.05	0.833	364504	DD_cage
3.3	1.36	1.23	94.65	0.808	385364	DD_cage
4.15	1.67	1.47	95.76	0.838	413692	DD_cage
4.85	1.88	1.74	95.61	0.847	400632	DD_cage
5.4	1.78	2.09	95.7	0.87	400454	DD_cage
5.26	1.76	1.98	96.1	0.864	400054	D_cage
4.57	1.58	1.69	96.39	0.863	413040	D_cage
3.69	1.28	1.41	95.37	0.844	384882	D_cage
4.88	1.73	1.81	96.23	0.854	383264	D_cage
4.89	1.7	1.8	96.3	0.866	399378	D_cage
4.66	1.6	1.7	96.09	0.873	431524	D_cage
5.75	0.65	2.7	95.49	0.906	464346	wound
5.56	0.64	2.7	95.42	0.878	388378	wound