A Hybrid Methodology for Inventory Classification with the Application of Machine Learning Algorithms

A thesis submitted in partial fulfillment of the requirements for the degree

BACHELOR OF SCIENCE IN INDUSTRIAL & PRODUCTION ENGINEERING

Anik Roy
201736002

Ashiqur Rahman Khan
201736004

MD. Fahim Montasir
201736022

Under the guidance of

Dr. Kais Bin Zaman

Professor, Department of IPE, BUET

Department of Industrial and Production Engineering
Military Institute of Science and Technology (MIST)
Mirpur Cantonment, Mirpur, Dhaka-1216

Responsibilities of Industrial Engineer

- Reduce Waste
- Minimize Cost
- Minimize Production and Inventory Time
- Solving Process Issues
- Maximize Process Efficiency
- Boost Profit
- Optimize Production Process

SUPPLY CHAIN MANAGEMENT

"Supply chain strategies require a total systems view of the links in the chain that work together efficiently to create customer satisfaction at the end point of delivery to the consumer. As a consequence, costs must be lowered throughout the chain by driving out unnecessary expenses, movements, and handling. Efficiency must be increased, and bottlenecks removed. The supply-chain system must be responsive to customer requirements."

-By Tony Hines

Drivers of SUPPLY CHAIN

Logistical Drivers	Cross-Functional Drivers
Facilities	Information
Inventory	Sourcing
Transportation	Pricing

Why Inventory?

Objectives of Inventory Management

-To ensure a continuous supply of materials and stock so that production should not suffer at the time of customers demand,

-To avoid both overstocking and under-stocking of inventory,

-To optimize various costs indulged with inventories like purchase cost, carrying cost, storage cost,

-To ensure the quality of goods at reasonable prices,

-To maintain a systematic record of inventory,

-To facilitate furnishing of data for short and long-term planning with a controlled inventory.

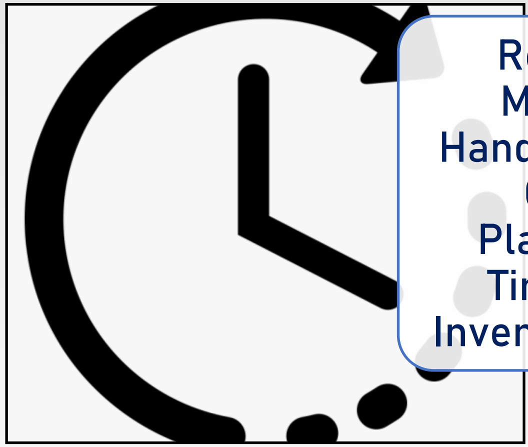
Our Focus

Inventory Item Classification

Cost

Minimizes Inventory Cost that associates reducing Production Cost

Time

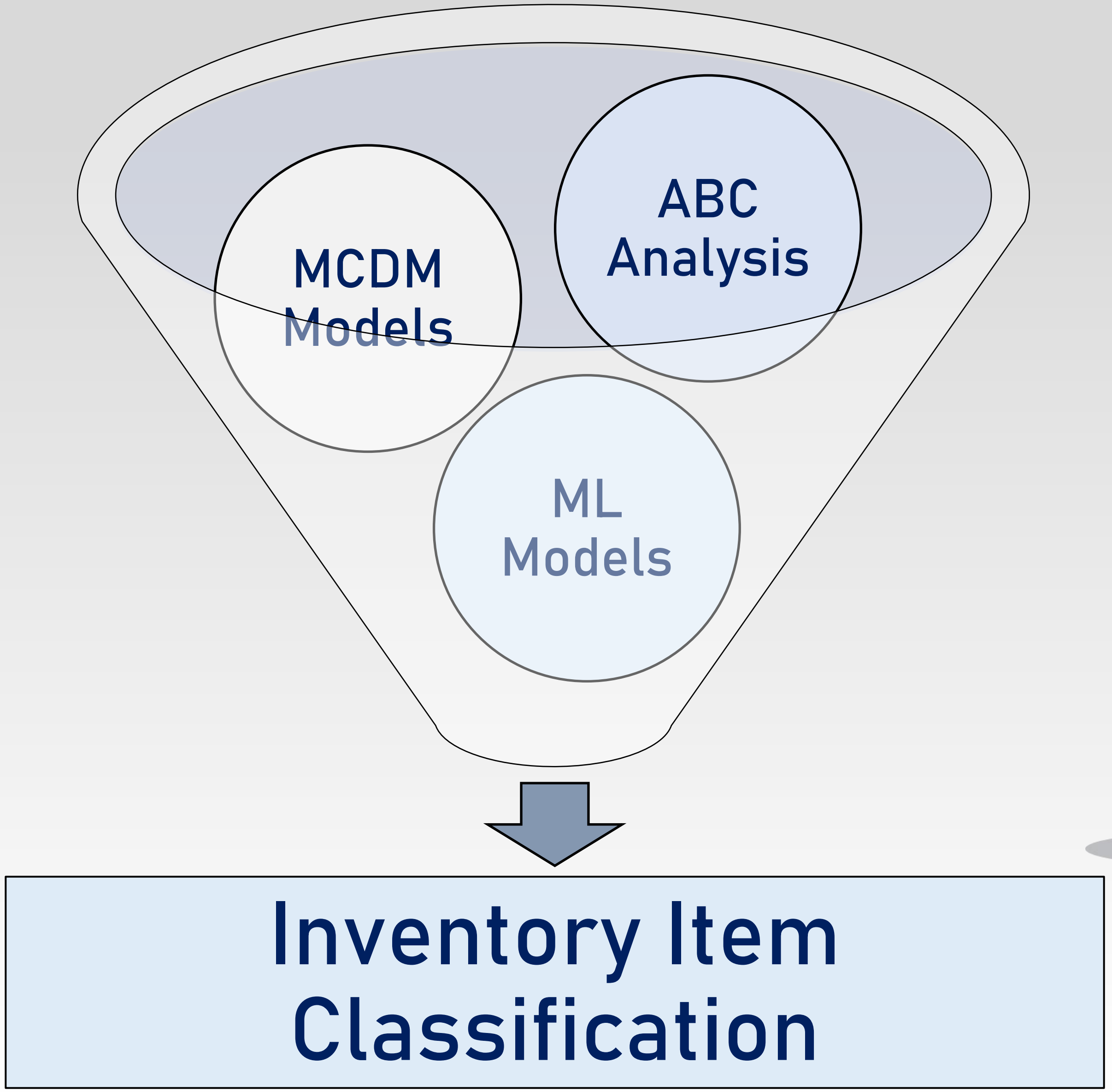


Reduces Material Handling Time, Order Placement Time, and Inventory Time.

Thesis Topic

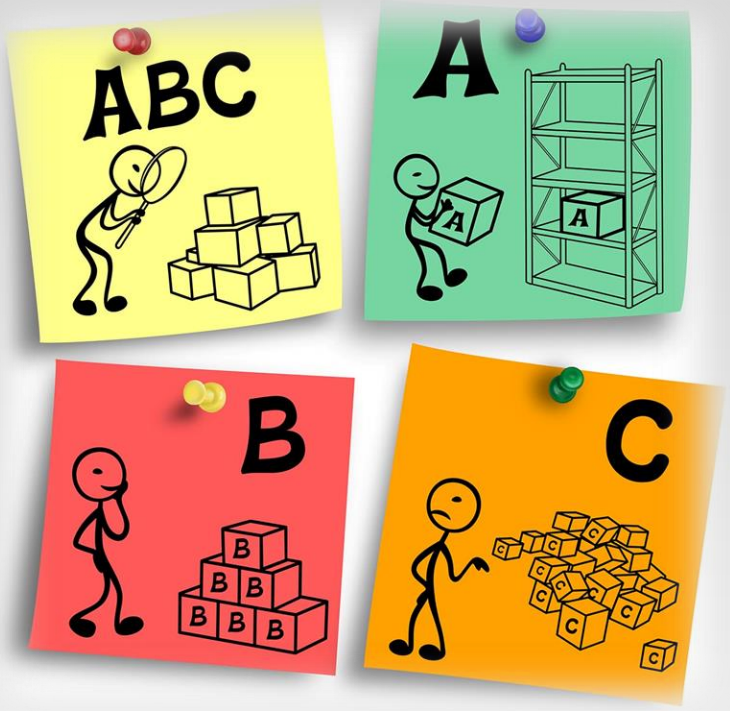
"A HYBRID METHODOLOGY FOR INVENTORY CLASSIFICATION WITH THE APPLICATION OF MACHINE LEARNING ALGORITHMS"

Classification Process



ABC Analysis Process

ABC Analysis basically classifies inventory items based on the annual consumption value (ACV) of the items. This analysis is done by 80/20 distribution or Pareto Principle.



A	B	C
'A' items – 20% of the items accounts for 70% of the ACV of items	'B' items – 30% of the items accounts for 25% of the ACV of items	'C' items – 50% of the items accounts for 5% of the ACV of items

Conventional ABC Analysis only accounts for Annual Consumption Value(ACV) of each item. But there are other major criteria which may affect the inventory classification.

Statistical Model Process

Multiple-Criteria Decision-Making (MCDM)

This method explicitly evaluates multiple conflicting criteria in decision making.

Multiple-Criteria

- Raw Data Set
- Conflicting Criteria

One New Criterion

- Utilizing different MCDM methods
- Weighted Scores
- Considering Beneficial and Non-beneficial Attributes

Reordering the Data Set

- Considering the newly generated criterion
- Ascending order or Descending order

ABC Analysis

- Following Pareto Principle
- Done for different MCDM methods
- For existing Inventory-items.

Here, MCDM method is used for Classification purpose.

Multiple-Criteria Decision-Making (MCDM)

There are several MCDM methods available. Such as-

- ❑Analytical Hierarchy Process(AHP)
- ❑Analytic Network Process (ANP)
- ❑Best Worst Method (BWM)
- ❑Simple Multi-Attribute Rating Technique (SMART)
- ❑Stratified Multi Criteria Decision Making (SMCDM)
- ❑Simple Additive Weight (SAW)
- ❑Weighted product model (WPM)
- ❑Weighted sum model (WSM)
- ❑VIKOR method
- ❑Technique for the Order of Prioritization by Similarity to Ideal Solution (TOPSIS)

After analyzing various research articles and implementation processes, following three MCDM method were selected for ABC Classification purpose-

- ❑Analytical Hierarchy Process(AHP)
- ❑Simple Additive Weight (SAW)
- ❑VIKOR method

What about the Newly added items to the inventory?

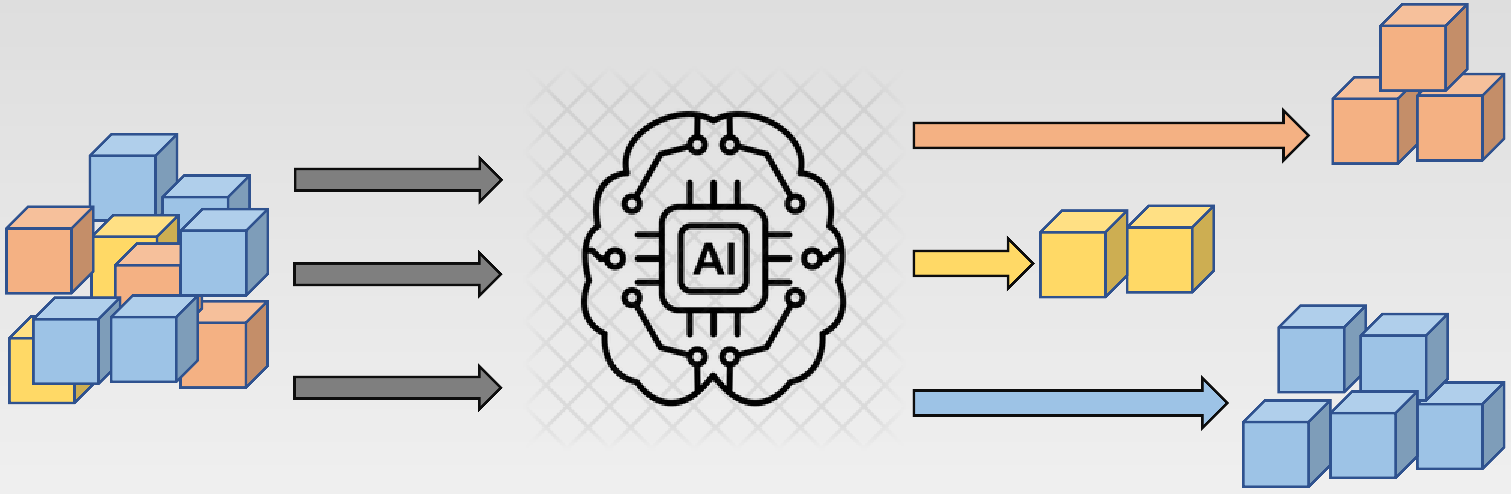
Should the whole previous activities be repeated for classifying the newly added item?

Reasons Behind this-

- Time Consuming
- Less Efficient
- Malfunctioning
- Re-classification of existing items
- Confused Classification

What would be done to solve this issue?

Machine Learning Algorithms can be implemented to solve this issue.



Data about the newly added item to the inventory

Trained with existing item- classification Machine Learning Model

Predicted Class for New Inventory Items

Machine Learning Model

Supervised Learning	Unsupervised Learning
Support Vector Machining (SVM)	K-means
Artificial Neural Network (ANN)	AdaBoost
Gaussian Process Classifier (GPC)	Apriori algorithm
K-Nearest Neighbors (KNN)	Principal Component Analysis (PCA)

Here, Supervised Machine Learning Algorithms are used that the class can be predicted precisely for newly added inventory item with the help of previously predicted data classes.

History on Hybrid Methodology for Multi-Attribute Inventory Analysis

Machine Learning Algorithms + MCDM Methods = Hybrid Methodology

10 June 2016, Hasan Kartal et al published "An integrated decision analytic framework of machine learning with multi-criteria decision making for multi-attribute inventory classification"

Inventory Classification	MCDM Methods	Machine Learning Algorithms
▪ ABC Analysis	▪ SAW ▪ AHP ▪ VIKOR	▪ ANN ▪ SVM, ▪ Naïve Bayes, ▪ Bayesian network

History on Hybrid Methodology for Multi-Attribute Inventory Analysis

Kartal et al (2013)	Ramanathan et al (2006)	Davood Sabaei et al (2015)	Chung-Hsing et al (2002)	Ali Jahan et al (2011)
Support Vector Machines for Multi-Attribute ABC Analysis	A weighted linear optimization for inventory classification	A review of multi-criteria decision-making methods for enhanced maintenance delivery	A problem-based selection of multi-attribute decision-making methods	A comprehensive VIKOR method for material selection
SVM is highly applicable in the inventory settings	ABC inventory classification with multiple-criteria using weighted linear optimization	MCDM techniques, comparison analysis among different MCDM methods	Different datasets may result in different classes for a particular inventory	Methodology on updated VIKOR method with different examples

Our Proposed Model

Selection of MCDM Models

SAW

- Basic and mostly used MCDM model to evaluate any multi-attributes inventory analysis
- Works with maximum and minimum value of a particular attribute

AHP

- Comparative analysis among the attributes
- Consistency of the attributes

VIKOR

- Optimized MCDM model
- A compromised solution based on the mutual concessions

Simple Additive Weight (SAW)

Linear Normalization, $k_{ij} = \begin{cases} \dfrac{x_{ij}}{max_i\, x_{ij}}, & if\ j\ is\ a\ beneficial\ attribute \\ \dfrac{min_i\, x_{ij}}{x_{ij}}, & if\ j\ is\ a\ non-beneficial\ attribute \end{cases}$

i =1, 2……., m; j =1, 2…. n.

Performance, $\mathrm{V_i} = \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{w_j k_{ij}}$; j =1, 2,…m.

Analytical Hierarchy Process(AHP)

Saaty's Scale for AHP

Rank	Description
1.00	Equally Important
3.00	Moderately Important
5.00	Strongly Important
7.00	Significantly Important
9.00	Extremely Important

Pair-Wise Matrix by AHP

	mfg	sell	demand	fcost	rma
mfg	1	1/7	1/5	2	1/3
sell	7	1	2	9	3
demand	5	1/2	1	7	3
fcost	1/2	1/9	1/7	1	1/5
rma	3	1/3	1/3	5	1

Analytical Hierarchy Process(AHP)

Criteria Weight Determination by AHP

	mfg	sell	demand	fcost	rma	criteria_weight
mfg	0.060606	0.068519	0.054407	0.083333	0.044205	0.062214
sell	0.424242	0.479157	0.544070	0.375000	0.398248	0.444143
demand	0.303030	0.239578	0.272035	0.291667	0.398248	0.300912
fcost	0.030303	0.053186	0.038901	0.041667	0.026550	0.038121
rma	0.181818	0.159559	0.090588	0.208333	0.132749	0.154610

VIšekriterijumsko KOmpromisno Rangiranje (VIKOR)

$$w_j = \sum_{i=1}^{n} w_i f_{ij}$$

Where,
f_(ij) = value of jth alternative and ith criterion
w_i = weight of ith criterion

$$S_j = \sum_{i=1}^{n} w_i (f_i^* - f_{ij})/(f_i^* - f_i^-)$$

$$R_j = \max_i [w_j (f_i^* - f_{ij})/(f_i^* - f_i^-)]$$

$$Q_j = \frac{v(S_j - S^*)}{(S^- - S^*)} + (1 - v)\frac{(R_j - R^*)}{(R^- - R^*)}$$

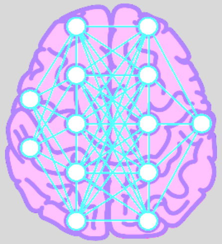
Selection of ML Algorithms

ANN
- Works well with non-linear data

GPC
- Higher accuracy with higher training time

KNN
- Comparatively Higher accuracy with lower training time

SVM
- Mostly used on classification models



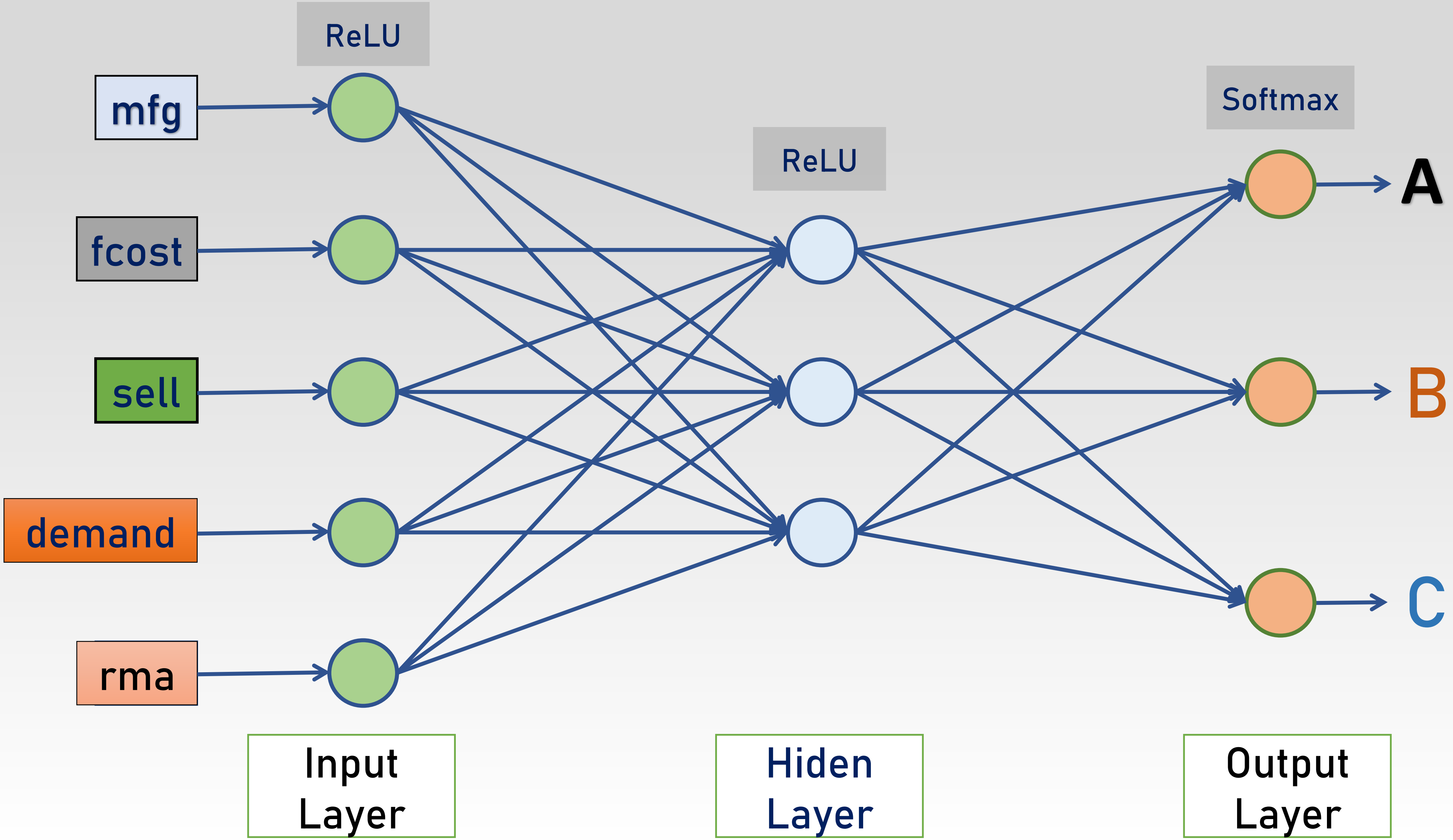
ARTIFICIAL NEURAL NETWORK

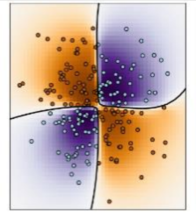
Supervised Machine Learning Method

Three Components

- **Input** — Takes extraordinary values and weights
- **Transfer Function** — Translates the input signals to output signals
- **Output** — Calculated for the known class

ARTIFICIAL NEURAL NETWORK

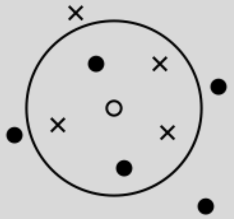




GAUSSIAN PROCESS CLASSIFICATION

- Nonparametric classification method
- Assumes some prior distribution on the underlying probability densities
- Confidence Interval can be developed from the probability densities
- Lazy Learning Algorithm

K-Nearest Neighbor



- A simple, supervised machine learning algorithm

- Solve both classification and regression problems

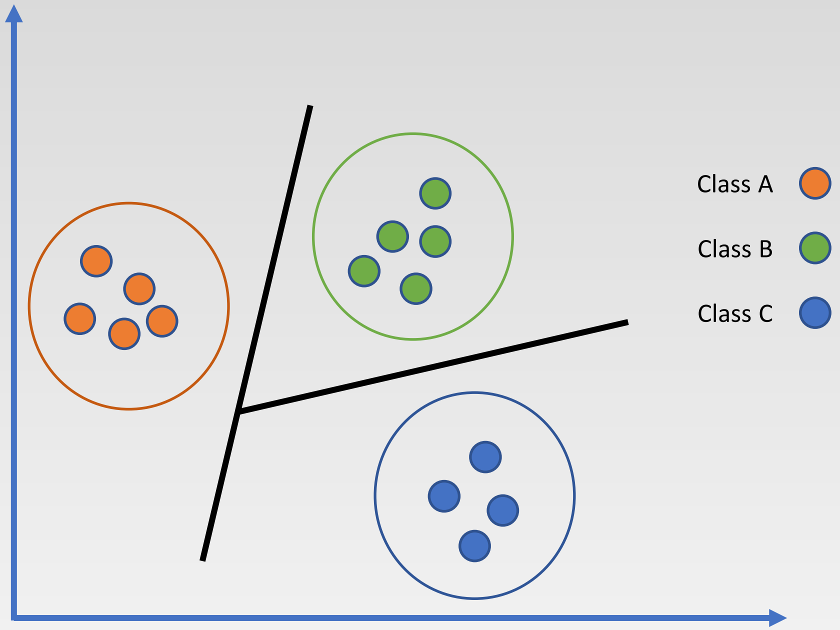
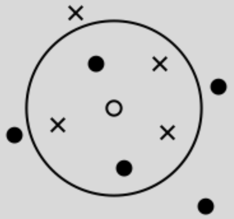


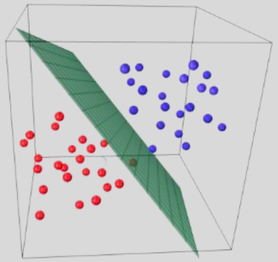
Fig: Visual representation of K-Nearest Neighbor

- The similarities of data points are measured by the Euclidean distance from a point to point



K-Nearest Neighbor (KNN)

- Calculate the distance between each row of training and testing data
- Sort the calculated distances in ascending order based on distance values
- Get top k rows from the sorted array
- Get the most frequent class of these rows
- Return the predicted class for test data

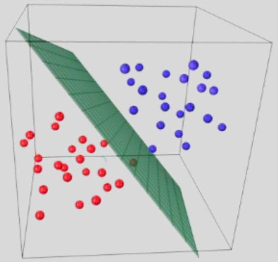


SUPPORT VECTOR MACHINE

- Support Vector Classifier (SVC) is the codomain of Deals with two class or multi class classification

- Maintains a maximum margin strategy that transformed into solving a complex quadratic programming problem

- Classifies data by finding the best hyper plane that separates all data points by maximum margin



SUPPORT VECTOR MACHINE

Maximum-margin hyper plane and margins for an SVM trained with samples from two classes

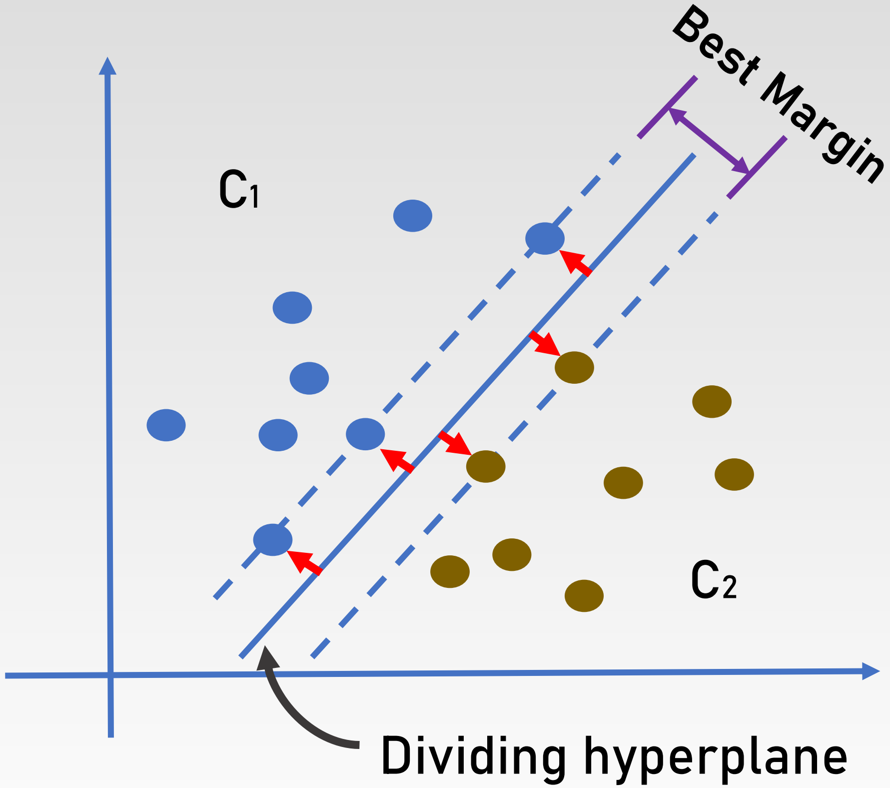


Fig: Visual representation of Support Vector Machine

Additional Approach to This Research

Newly Implemented Machine Learning Models for Inventory Item Classification:

Gaussian Process Classifier (GPC)

K-Nearest Neighbors (KNN)

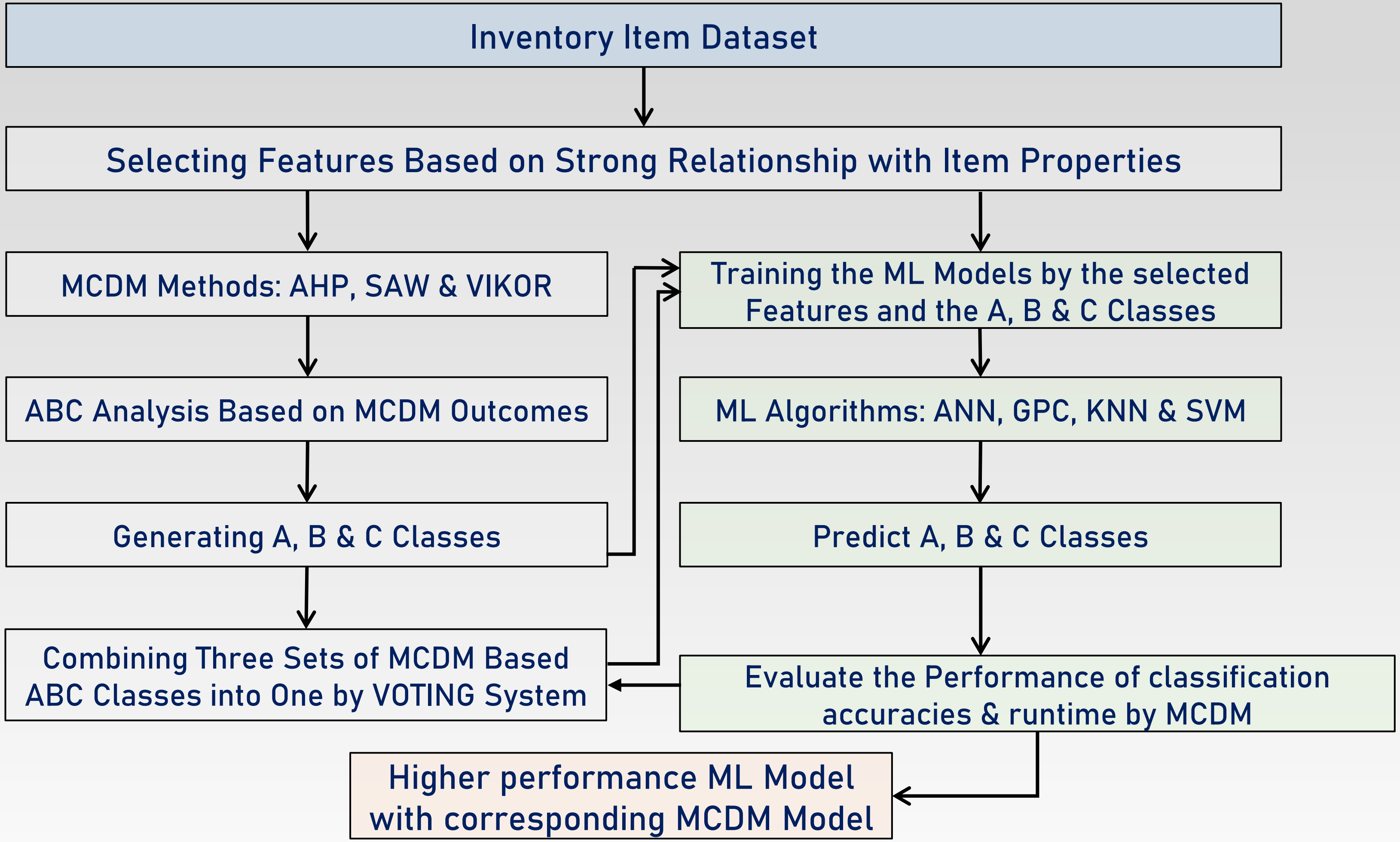
Additional Approach to This Research

“Voting System” for Inventory Item Classification:

Combines all MCDM method results into one.

Gives a Conclusive Classification for existing inventory items.

Methodology



Data Description

Data Source- Randomly generated dummy data

Selected Features for this Study

Features	Data Type	Dataset Type	Features Category
Manufacturing Cost	int64	Discrete	Non-beneficial
Fixed Cost	float64	Continuous	Non-beneficial
Selling Price	float64	Continuous	Beneficial
Demand per Week	int64	Discrete	Beneficial
Raw Material Availability	int64	Categorical (scale of 1 to 3)	Beneficial

Raw Material Availability: High = 3, Medium = 2 & Low = 1

Data Description

1.Manufacturing Cost per Product:

Manufacturing Cost = np.random.randint(50,500, n)

Where, n = number of items or rows

2.Fixed Cost per Product:

Fixed Cost = (manufacturing cost / 2) + (20% of manufacturing cost)

3.Selling Price:

Selling Price = (manufacturing cost + fixed cost) + (manufacturing cost + fixed cost) / 2

4.Demand per week:

Demand per Week = np.random.randint(1500, 7000, n)

5.Raw Material Availability:

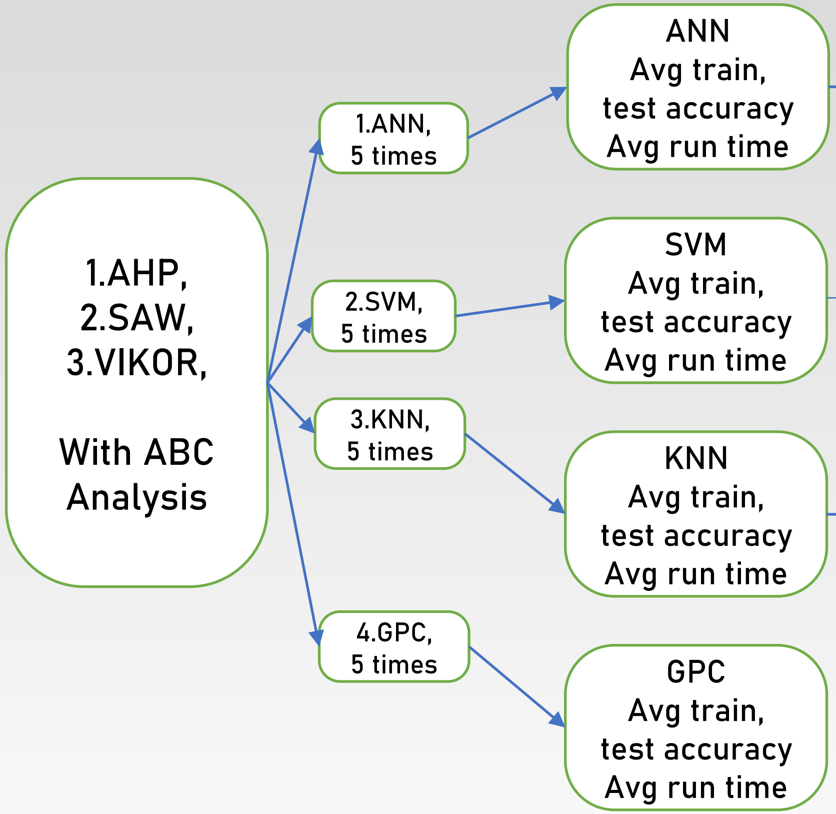
Raw material availability of a product is a categorical data

Generated Dataset 01 (100000 x 5) & Dataset 02 (150000 x 5) without any Duplication of Rows

Machine Specifications:

Machine: Virtual Machine powered by Kaggle
Processor: 4 Core
Clock Speed : 2.2 GHz
Ram: 16 GB

ABC Analysis of MCDM Methods and the Application of Machine Learning Algorithms



AHP 150K	train_Acc	test_Acc	time	performance
ANN	99.8	99.8	24.99	0.991
SVM	77.94	77.83	278.32	0.043
KNN	87.38	75.67	0.18	0.392
GPC	100	74.49	301.35	0.633

AHP 100K	train_Acc	test_Acc	time	performance
ANN	99.73	99.73	16.34	0.982
SVM	77.85	77.83	116.98	0.045
KNN	87.07	75.03	0.11	0.381
GPC	100	73.5	103.53	0.654

SAW 150K	train_Acc	test_Acc	time	performance
ANN	99.79	99.79	23.46	0.995
SVM	64.74	64.95	566.32	0.027
KNN	80.82	63.94	0.18	0.443
GPC	100	63.03	308.44	0.592

SAW 100K	train_Acc	test_Acc	time	performance
ANN	99.77	99.77	16.47	0.984
SVM	64.6	65.1	219.86	0.027
KNN	80.4	63.21	0.11	0.436
GPC	100	62.32	100.7	0.615

VIKOR150K	train_Acc	test_Acc	time	performance
ANN	98.26	98.26	13.56	0.925
GPC	100	73.42	113.49	0.701
SVM	91.93	92.04	117.51	0.335
KNN	88.07	77.2	0.11	0

VIKOR100K	train_Acc	test_Acc	Time	Performance
ANN	98.89	98.89	36.14	0.980
GPC	100	76.05	314.4	0.821
KNN	88.37	77.73	0.54	0.513
SVM	75.6	75.57	784.86	0

Discussion on Voting System

Item_Number	AHP_Class	SAW_Class	VIKOR_Class	Combo_Class
1	C	C	C	C
2	B	A	A	A
3	A	A	A	A
4	A	A	A	A
5	C	B	A	B

1. AHP based ABC Classes (AHP_Class)

2. SAW based ABC Classes (SAW_Class)

3. VIKOR based ABC Classes (VIKOR_Class)

For Item_Number 2:

AHP_Class	SAW_Class	VIKOR_Class
B	A	A

Combo_Class
A

Discussion on Voting System

Table with Conflicting Class

Item_Number	AHP_Class	SAW_Class	VIKOR_Class	Combo_Class
1	C	C	C	C
2	B	A	A	A
3	A	A	A	A
4	A	A	A	A
5	C	B	A	ABC

Table After Solving Conflict

Item_Number	AHP_Class	SAW_Class	VIKOR_Class	Combo_Class
1	C	C	C	C
2	B	A	A	A
3	A	A	A	A
4	A	A	A	A
5	C	B	A	B

Table with Conflicting Class

Class	Count	Percentage
C	76227	76.227%
ABC	11252	11.252%
B	7438	7.438%
A	5083	5.083%
Total	100000	100%

Solution For Item_Number 5:

AHP_Class	SAW_Class	VIKOR_Class
• C	• B	• A

Combo_Class
• B

Table After Solving Conflict

Class	Count	Percentage
C	76227	76.227%
B	13543	13.543%
A	10230	10.230%
Total	100000	100%

Nearly Expected Percentage of Classes

ABC Analysis of MCDM & Combined Approach Methods and the Application of Machine Learning Algorithms

1.AHP,
2.SAW,
3.VIKOR,
4.Combo

With ABC Analysis

1.ANN, 5 times

2.SVM, 5 times

3.KNN, 5 times

4.GPC, 5 times

ANN
Avg train, test accuracy
Avg run time

SVM
Avg train, test accuracy
Avg run time

KNN
Avg train, test accuracy
Avg run time

GPC
Avg train, test accuracy
Avg run time

AHP 150K	train_Acc	test_Acc	time	performance
ANN	99.8	99.8	24.99	0.991
SVM	77.94	77.83	278.32	0.043
KNN	87.38	75.67	0.18	0.392
GPC	100	74.49	301.35	0.633

AHP 100K	train_Acc	test_Acc	time	performance
ANN	99.73	99.73	16.34	0.982
SVM	77.85	77.83	116.98	0.045
KNN	87.07	75.03	0.11	0.381
GPC	100	73.5	103.53	0.654

SAW 150K	train_Acc	test_Acc	time	performance
ANN	99.79	99.79	23.46	0.995
SVM	64.74	64.95	566.32	0.027
KNN	80.82	63.94	0.18	0.443
GPC	100	63.03	308.44	0.592

SAW 100K	train_Acc	test_Acc	time	performance
ANN	99.77	99.77	16.47	0.984
SVM	64.6	65.1	219.86	0.027
KNN	80.4	63.21	0.11	0.436
GPC	100	62.32	100.7	0.615

VIKOR150K	train_Acc	test_Acc	time	performance
ANN	98.26	98.26	13.56	0.925
GPC	100	73.42	113.49	0.701
SVM	91.93	92.04	117.51	0.335
KNN	88.07	77.2	0.11	0

VIKOR100K	train_Acc	test_Acc	time	performance
ANN	98.89	98.89	36.14	0.980
GPC	100	76.05	314.4	0.821
KNN	88.37	77.73	0.54	0.513
SVM	75.6	75.57	784.86	0

COMBO 150K (SAW)	train_Acc	test_Acc	time	performance
ANN	99	99	20.45	0.97
SVM	77	77	406.12	0.01
KNN	87	76	0.16	0.42
GPC	100	77	325.12	0.55

COMBO 100K (SAW)	train_Acc	test_Acc	time	performance
ANN	98	98	18.61	0.94
SVM	77	78	206.14	0.04
KNN	87	76	0.11	0.43
GPC	100	75	121.67	0.58

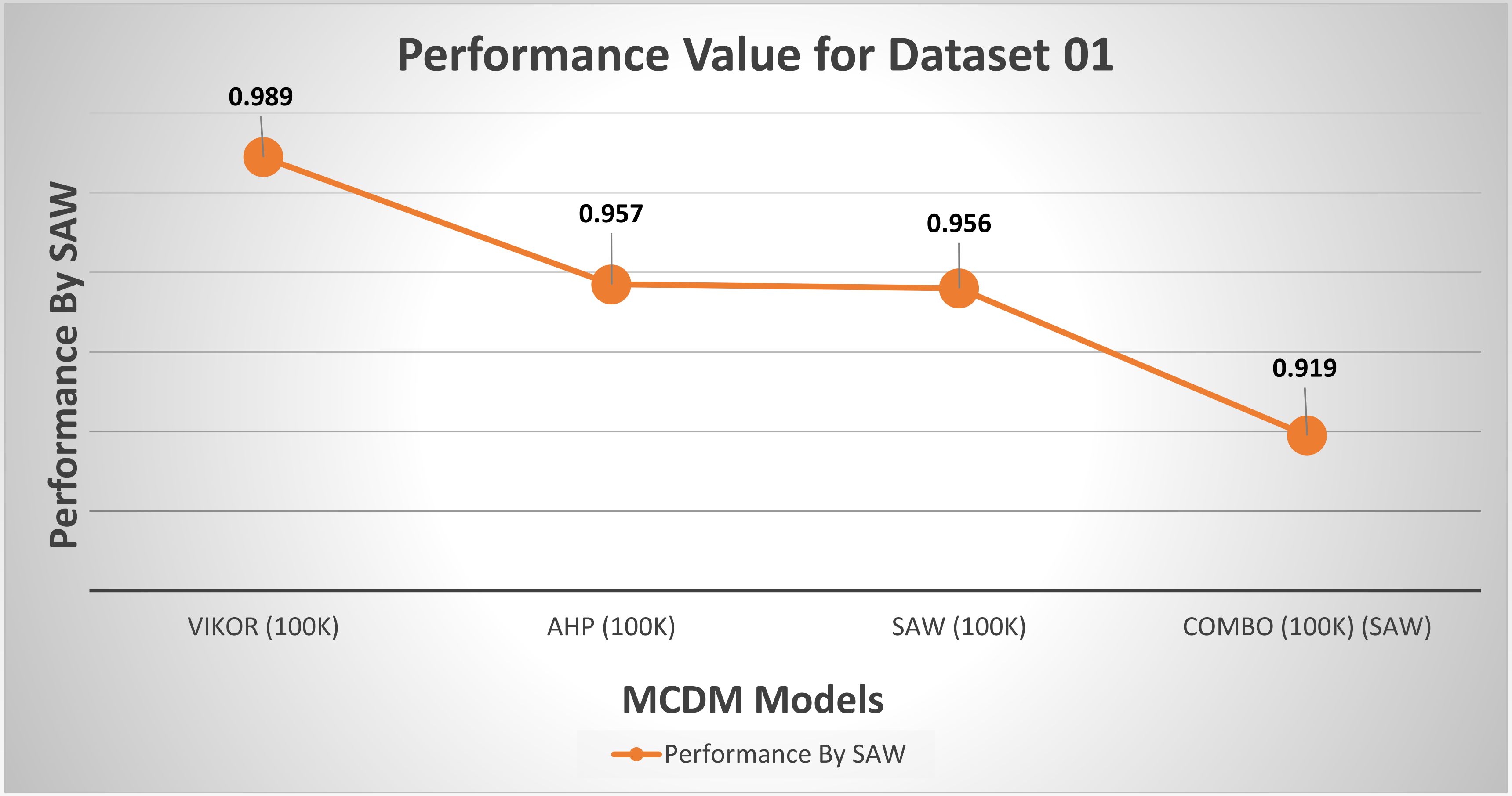
RESULT

Overall Performance Measure by SAW for Dataset 01

Dataset 01 (100k)	MCDM	ML	TRAIN ACCURACY	RUN TIME	INDIVIDUAL PERFORMANCE	Performance By SAW
0	VIKOR (100K)	ANN	98.26	13.56	0.980	0.989
1	AHP (100K)	ANN	99.73	16.34	0.982	0.957
2	SAW (100K)	ANN	99.77	16.47	0.984	0.956
3	Combo (100k) (SAW)	ANN	98	18.61	0.944	0.919

RESULT

Performance Value Vs MCDM Models for Dataset 01



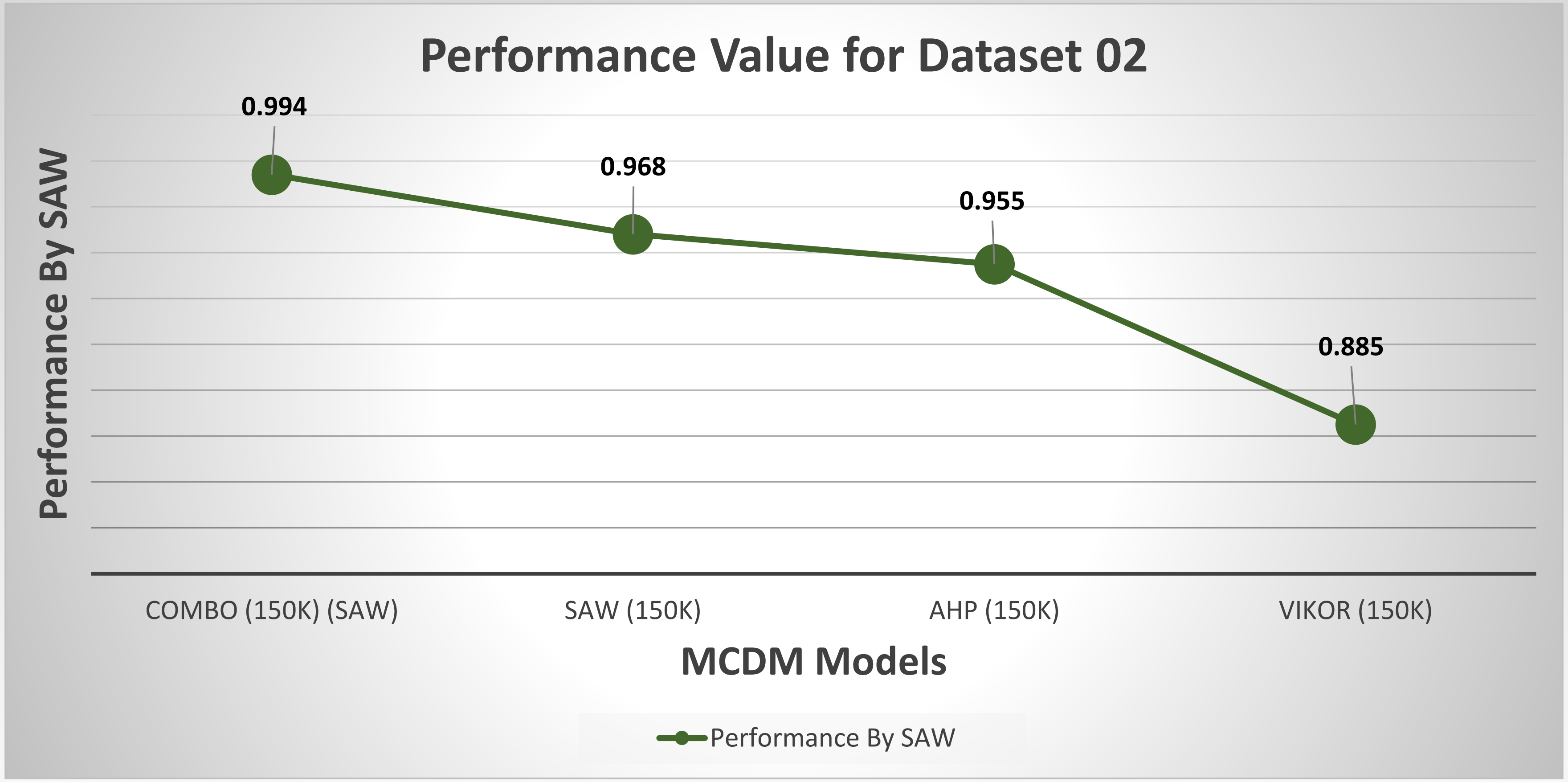
RESULT

Overall Performance Measure by SAW for Dataset 02

Dataset 02 (150k)	MCDM	ML	TRAIN ACCURACY	RUN TIME	INDIVIDUAL PERFORMANCE	Performance By SAW
0	Combo (150k) (SAW)	ANN	99	20.45	0.973	0.994
1	SAW (150k)	ANN	99.79	23.46	0.995	0.968
2	AHP (150K)	ANN	99.8	24.99	0.991	0.955
3	VIKOR (150K)	ANN	98.89	36.14	0.925	0.885

RESULT

Performance Value Vs MCDM Models for Dataset 02



Conclusion

1. Artificial Neural Network (ANN) is found the most balanced machine learning algorithm among the four ML algorithms.

2. VIšekriterijumsko KOmpromisno Rangiranje (VIKOR) is found the highest performing MCDM model for small dimensional dataset.

3. Combined Approach of SAW, AHP and VIKOR shows better performance than individual MCDM models for larger dimensional dataset.

FUTURE WORK

Better research can be done if real-life data is available

More accurate ML algorithms may be introduced in future

Different MCDM methods may be applied

THANK YOU!