IMPACT OF ARTIFICIAL INTELLIGENCE ON INDUSTRIES

UNIVERSITY OF MUMBAI

DEPARTMENT OF STATISTICS
VIDYANAGARI, MUMBAI-400098

CERTIFICATE

This is to certify that students of M.SC. Part 2, have successfully completed the project entitled **- "IMPACT OF ARTIFICIAL INTELLIGENCE ON INDUSTRIES"**

during the Academic Year 2023-2024.

The Team Comprises of:-

- **ABHISHEK KASHID**
- **GAURAV HUMANE**
- **HARSH MANDARE**
- **SAURABH TRIPATHI**
- **VINAYAK HAPTE**

This work is the best of our knowledge and belief is original.

Dr. V.U. DIXIT
(Associate professor & Project Guide)

Dr. SANTOSH P. GITE
(Head of the department)

ACKNOWLEDGEMENT

From the conception of topic to the final presentation ,we are extremely thankful to our guide **Dr. Vaijayanti U. Dixit** (Assistant professor of department of statistics at University of Mumbai) for her immense support and guidance throughout the process of making the entire project . It was only with the help of her suggestions and corrections , at every step that we were able to complete the project on time and achieve our objectives to the satisfaction.

We are extremely thankful to Dr. Alok. D. Dabade and Professor Sheetal Chabukeshwar for their support and guidance and providing all the necessary information. Also, we would like to thank Dr. Santosh Gite(Head of the department) for providing us the opportunity to do the project work and giving us all the support and guidance , which made us complete the project duly.

Last but not least , we would like to give our special thanks to our respondents who diligently and honestly filled our survey purely for our academy purposes and thanking the time out of their schedule to answers our questions genuinely.

INDEX

INTRODUCTION

AI, or artificial intelligence, is the development of computer systems that can perform tasks that typically require human intelligence, such as learning, problem-solving, and decision- making. It's like giving computers a brain to think and act like people do. AI uses data and rules to figure out patterns and make predictions or decisions. For example, AI can help with things like recognizing faces in photos, recommending products you might like, or playing games. AI is all around us in smartphones, smart speakers, self-driving cars, and many other technologies that make our lives easier and more convenient. It is a rapidly evolving field that is transforming various industries and aspects of our daily lives, from virtual assistants like Siri and Alexa to self-driving cars and personalized recommendations on streaming platforms.

- **History of AI**

The history of AI dates back to the mid-20th century when researchers began exploring the concept of artificial intelligence. Early developments include Alan Turing's work on computing and the Turing Test in the 1950s, followed by the development of expert systems in the 1970s and neural networks in the 1980s. The field experienced periods of both excitement and disappointment, known as "AI summers" and "AI winters," as progress fluctuated. Breakthroughs in machine learning, particularly deep learning, have fueled significant advancements in recent years, leading to applications in various domains like image recognition, natural language processing, and robotics. Today, AI is a rapidly evolving field with growing impacts on society, from autonomous vehicles to healthcare.

- Impacts of AI on Industries

Artificial Intelligence (AI) has had a transformative impact on various industries, revolutionizing the way businesses operate and deliver products and services and in sectors too.

- AI in Finance

AI in personal finance applications, such as Intuit Mint or TurboTax, is disrupting financial institutions. Applications such as these collect personal data and provide financial advice. Other programs, such as IBM Watson, have been applied to the process of buying a home. Today, artificial intelligence software performs much of the trading on Wall Street.

- AI in Software Coding and IT Processes

New generative AI tools can be used to produce application code based on natural language prompts, but it is early days for these tools and unlikely they will replace software engineers soon. AI is also being used to automate many IT processes, including data entry, fraud detection, customer service, and predictive maintenance and security.

- AI in Manufacturing

Manufacturing has been at the forefront of incorporating robots into the workflow. For example, the industrial robots that were at one time programmed to perform single tasks and separated from human workers, increasingly function as cobots: Smaller, multitasking robots that collaborate with humans and take on responsibility for more parts of the job in warehouses, factory floors and other workspaces.

- AI in Transportation

In addition to AI's fundamental role in operating autonomous vehicles, AI technologies are used in transportation to manage traffic, predict flight delays, and make ocean shipping safer and more efficient. In supply chains, AI is replacing traditional methods of forecasting demand and predicting disruptions, a trend accelerated by COVID-19 when many companies were caught off guard by the effects of a global pandemic on the supply and demand of goods.

- AI in Banking

Banks are successfully employing chatbots to make their customers aware of services and offerings and to handle transactions that don't require human intervention. AI virtual assistants are used to improve and cut the costs of compliance with banking regulations. Banking organizations use AI to improve their decision-making for loans, set credit limits and identify investment opportunities.

- AI in Healthcare

The biggest bets are on improving patient outcomes and reducing costs. Companies are applying machine learning to make better and faster medical diagnoses than humans. One of the best-known healthcare technologies is IBM Watson. It understands natural language and can respond to questions asked of it. The system mines patient data and other available data sources to form a hypothesis, which it then presents with a confidence scoring schema. Other AI applications include using online virtual health assistants and chatbots to help patients and healthcare customers find medical information, schedule appointments, understand the billing process and complete other administrative processes. An array of AI technologies is also being used to predict, fight and understand pandemics such as COVID- 19.

Advantages of AI	Disadvantages of AI
• Efficiency: AI can perform tasks faster and with fewer errors than humans.	• Job Displacement: AI automation may lead to job loss in certain sectors.
• Automation: It can automate repetitive tasks, freeing up time for more complex work.	• Dependence: Overreliance on AI systems could lead to loss of critical skills in humans.
• Cost Savings: Implementing AI can lead to cost reductions in various industries	• Regulatory Challenges: There may be difficulties in regulating AI technologies to ensure they are used responsibly and ethically.
• Safety: AI can be used in dangerous environments, reducing risks to human workers.	• Privacy Concerns: AI systems may collect and analyze personal data, raising privacy issues.
• Accuracy: AI algorithms can analyze vast amounts of data and make accurate predictions.	• Bias: AI algorithms can perpetuate biases present in the data they're trained on.
• Accessibility: AI-powered tools can make services more accessible to people with disabilities.	• Lack of Transparency: Some AI algorithms operate as black boxes, making it difficult to understand their decision-making process.

Objectives

Following are the objectives of our survey :

1) To study the impact of Artificial Intelligence on the Efficiency of the work in the companies/ corporate offices.

2) To analyze the impact of Artificial Intelligence on the Expenditure of the various companies.

3) To study the impact of Artificial Intelligence on Customer Satisfaction

Research Methodology

<u>Steps involved in conducting the survey:</u>

- Defining our objective and scope of the survey.
- Literature survey -Review existing studies and theories related to AI and its impact.
- Develop specific questions and hypotheses that our research will address.
- Pilot Survey.
- Modifying questionnaires.
- Data Collection (Used Primary data).
- Data coding and Data entry.
- Data analysis.
- Conclusion.
- Preparation of Project Report.

DATA COLLECTION

PILOT SURVEY

LIVE SURVEY

TWO MONTHS

292 RESPONSES

Methodology : We conducted a pilot survey of sample size 52. After which we made the necessary changes in our Questionnaire .We then surveyed 346 individually and via Google forms with the help of Non – probability Convenient Sampling. We cleaned the data after entering it in the MS-EXCEL. After data cleaning ,we were left with final sample size 292.

DATA CLEANING :

Removal of unwanted observations

Handling missing data

Managing unwanted outliers

Fixing structural errors

- ❖ Types of Questions
  - Dichotomous questions
  - Likert Scale questions
  - Categorical questions
  - Open ended questions

- ❖ Techniques used
  - Graphical Representation
  - Ordinal Logistic Regression
  - Confusion matrix
  - Likelihood ratio test
  - Ordinal Forest technique

- ❖ Software used
  - SPSS
  - R
  - MS Office

1) <u>Ordinal logistic regression</u>- Ordinal logistic regression is a statistical analysis method that can be used to model the relationship between an ordinal response variable and one or more explanatory variables. An ordinal variable is a categorical variable for which there is a clear ordering of the category levels. The explanatory variables may be either continuous or categorical. The response variable has a clear order, such as "low", "medium", and "high", but the intervals between these categories are not equal.

Ordinal logistic regression is an extension of logistic regression where the logit (i.e. the log odds) of a binary response is linearly related to the independent variables. If instead the response variable has k levels, then there are k-1 logits.Hence the output of an ordinal logistic regression will contain an intercept for each level of the response except one, and a single slope for each explanatory variable

Model- -

$$\Pi_i = \frac{e^{logit\, F_l}}{1 + e^{logit\, F_l}} \qquad , l = 1,2,..k-1$$

Logit function=logit F_l = $log\left(\frac{F_l}{1-F_l}\right) = log\left(\frac{\sum_{r=1}^{l} P_r}{1-\sum_{r=1}^{l} P_r}\right)$

$$= \beta_{0l} + \beta_{1l}x \quad ; l = 1,2,..k-1$$

where , β_{0l} $is\ the\ intercept\ for\ lth\ category\ of\ \Pi_i$

β_{1l} $is\ the\ regression\ coefficient\ associated\ with\ the\ predictor\ variablex$

<u>Assumptions-</u>

i. The dependent variable is measured on an ordinal level.
ii. One or more of the independent variables are either continuous, categorical or ordinal.
iii. No multi-collinearity - i.e. when two or more independent variables are highly correlated with each other.
iv. Proportional Odds - i.e. that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable. The odds ratios between different levels of the response variable are assumed to be constant.

2). Confusion matrix- A confusion matrix is a fundamental tool for evaluating the performance of classification models, particularly in binary and multiclass classification problems. It provides a detailed breakdown of the predicted and actual classifications made by the model.

In multiclass classification, the confusion matrix is extended to an n×n matrix where n is the number of classes. Each cell (i,j) in the matrix represents the number of instances of class i that were predicted as class j. This helps in understanding the performance of a classifier across multiple classes.

Structure of a Multiclass Confusion Matrix. For a problem with three classes (e.g., A, B, C), the confusion matrix will look like this:

Actual\Predicted	Predicted A	Predicted B	Predicted C
Actual A	A_{AA}	A_{AB}	A_{AC}
Actual B	B_{BA}	B_{BB}	B_{BC}
Actual C	C_{CA}	C_{CB}	C_{CC}

Where A_{AA} , B_{BB}, C_{CC} are the counts of correctly classified instances for each class, and the off-diagonal elements represent misclassifications.

Metrics Derived from the Multiclass Confusion Matrix

Several metrics can be calculated from the confusion matrix for each class and overall:

Accuracy: The overall proportion of correct predictions.

$$\text{Accuracy}=\frac{\sum_{i=1}^{n} A_{ii}}{\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}}$$

Precision for class i: The proportion of true positive predictions for class i among all predictions for class i

$$\text{Precision}_i =\frac{A_{ii}}{\sum_{j=1}^{n} A_{ji}}$$

Recall for Class i (Sensitivity or True Positive Rate): The Proportion of positive predictions for class i among all actual instances of class i.

$$\text{Recall}_i = \frac{A_{ii}}{\sum_{j=1}^{n} A_{ij}}$$

F1 Score for Class i: The harmonic mean of precision and recall for class i.

$$\text{F1 score}_i = 2 * \frac{Precision_i * Recall_i}{Precision_i + Recall_i}$$

3)Likelihood ratio test: Likelihood ratio test (LRT) typically includes several key pieces of information that help you understand whether the more complex model (full model) provides a significantly better fit to the data than the simpler model (reduced model).

Components of LRT

Log-Likelihood Values:

Log-Likelihood of the Reduced Model**:** Indicates the fit of the simpler model.

Log-Likelihood of the Full Model: Indicates the fit of the more complex model.

Likelihood Ratio Test Statistic (Deviance):

This is calculated as

$$G^2 = -2(\ln(\text{reduced model}) - (\text{full model}))$$

It measures the difference in fit between the two models.

Hypothesis:

H_0= Reduced model as good as full model

$H_{1=}$ Reduced model is not as good as full model

If P value >0.05, then we do not reject null hypothesis and conclude that

Reduced model is as good as full model

4)Ordinal Forest technique-The ordinal forest technique is a machine learning method specifically designed for ordinal regression tasks. It extends the traditional random forest algorithm to handle ordinal outcome variables, where the categories have a meaningful order but the distances between them are not necessarily equal.

The ordinal forest algorithm adapts the random forest technique for ordinal data by incorporating the order information into the tree-building process.

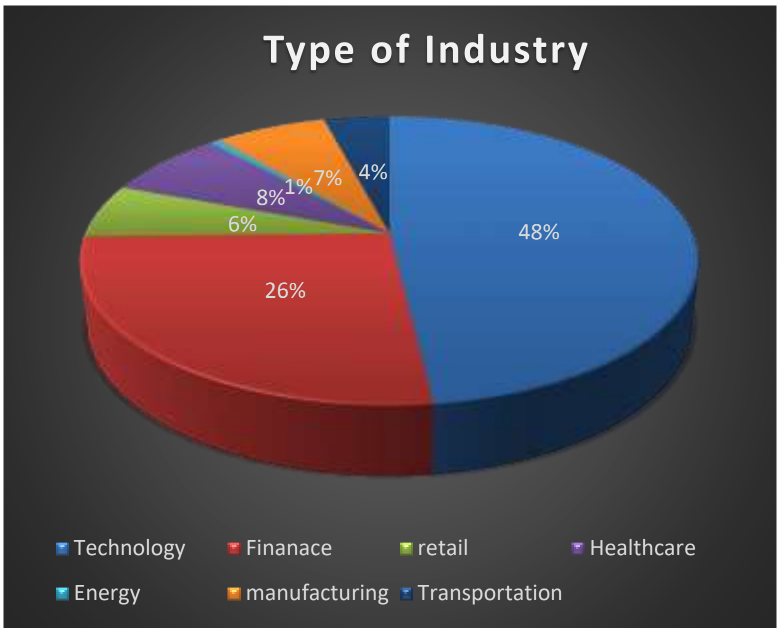
Here are some specific characteristics:

1. *Proportional Odds Assumption*:- Ordinal Forest does not rely on the proportional odds assumption, which is a common constraint in traditional ordinal regression models like the proportional odds logistic regression.
2. *Tree Construction*: - Trees are constructed using binary splits that respect the order of the outcome variable. This means the splits are chosen to maximize the separation of ordered categories while considering the ordinal nature of the data.
3. *Performance Measure*: - The algorithm optimizes splits based on performance measures suitable for ordinal data, such as the cumulative link loss function.
4. *Handling Multicollinearity*: - Similar to random forests, ordinal forests can handle multicollinearity and interaction effects among predictors.

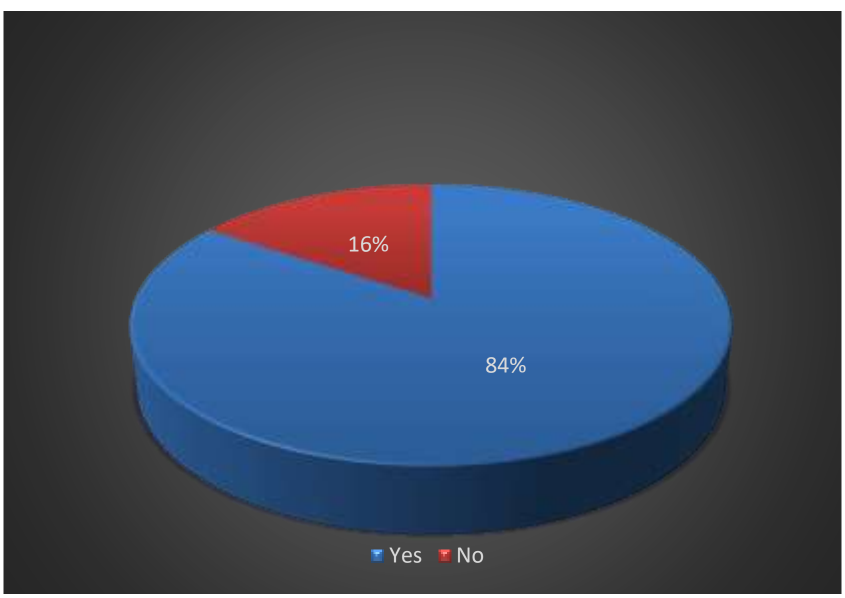
GRAPHICAL REPRESENTATION OF THE DATA:

The main purpose of graphical representation is to readily give some idea about the entire data and draw instant conclusions.

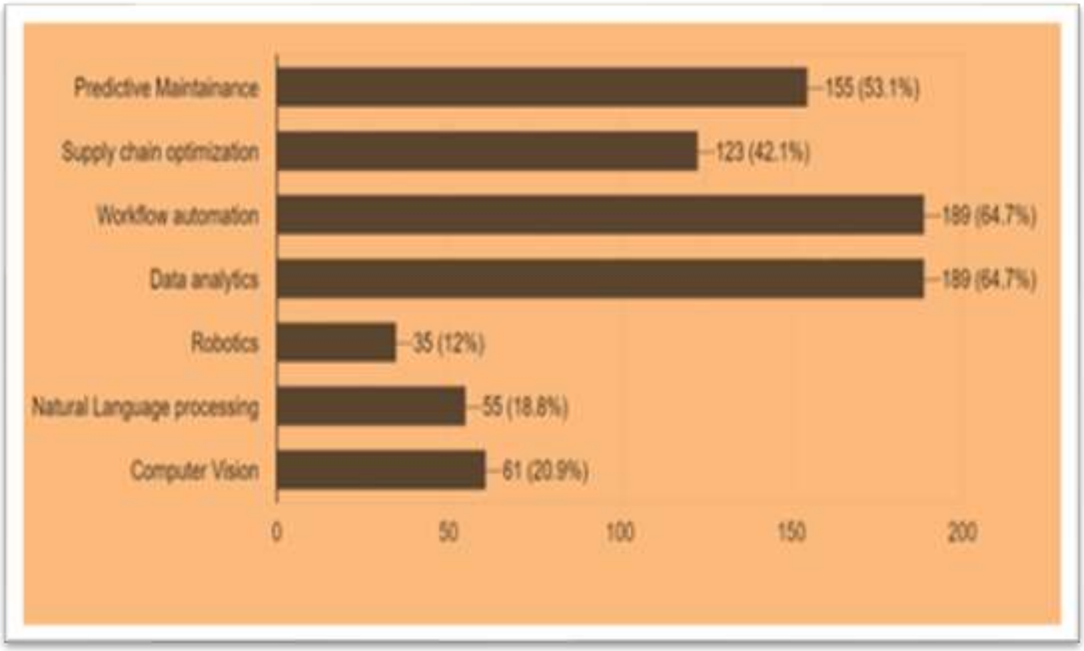
- Type of Industry



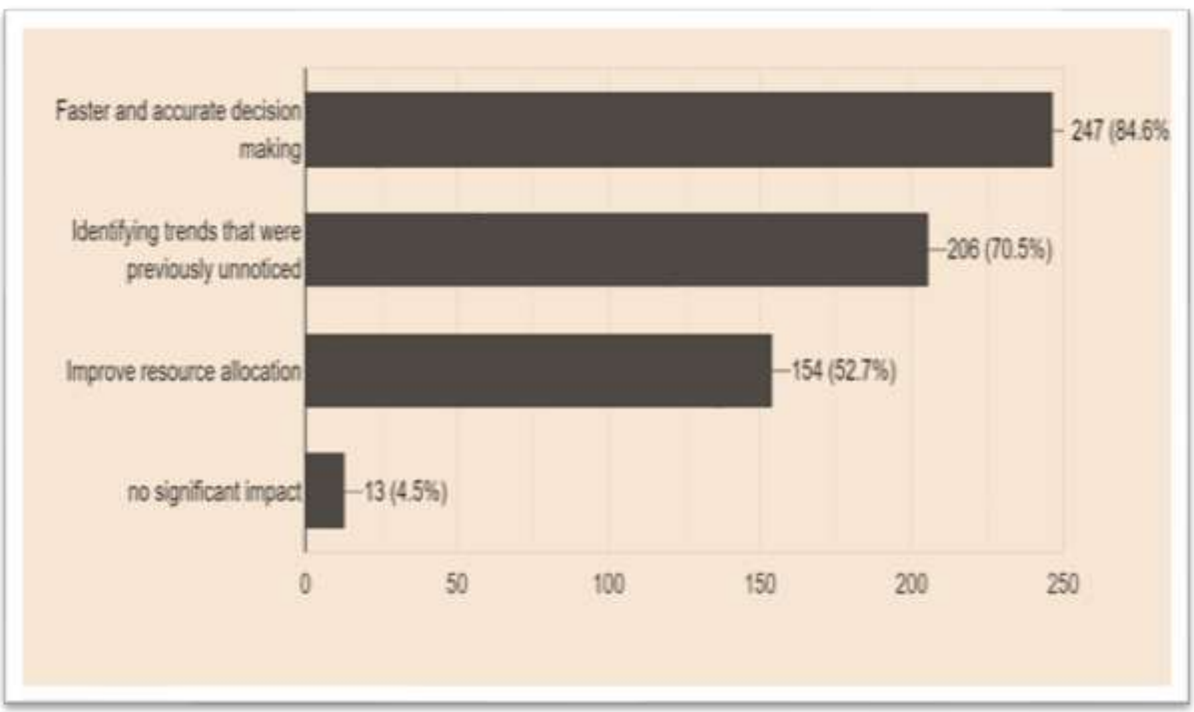
- The company currently utilize the Artificial intelligence



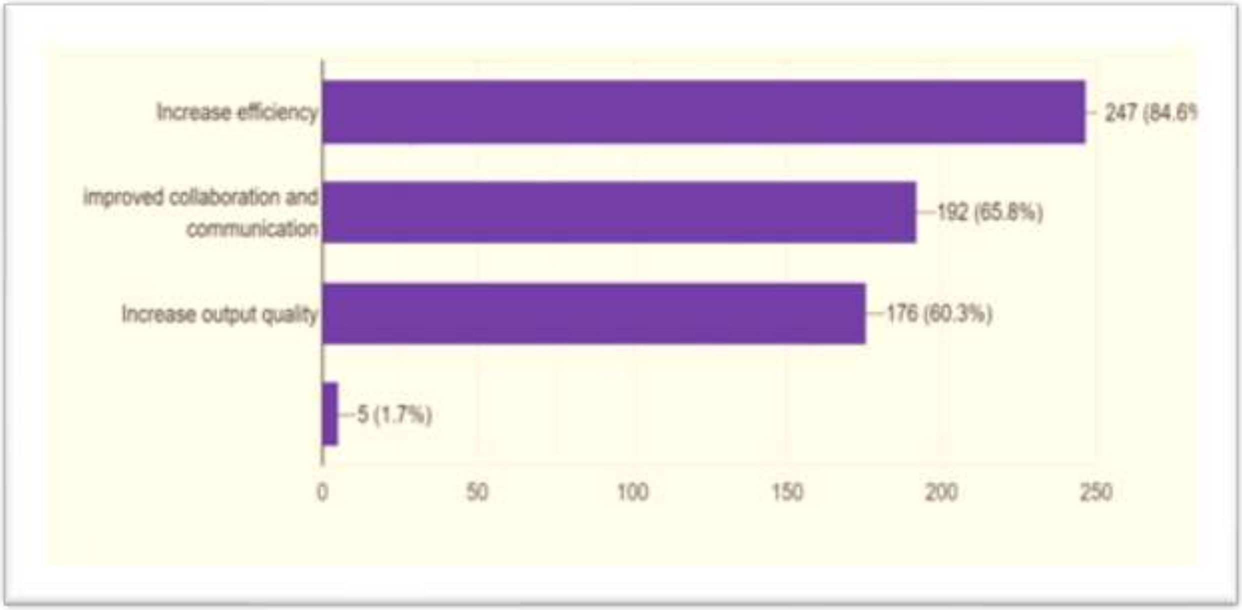
- AI applications significantly impact on the efficiency in your organization.



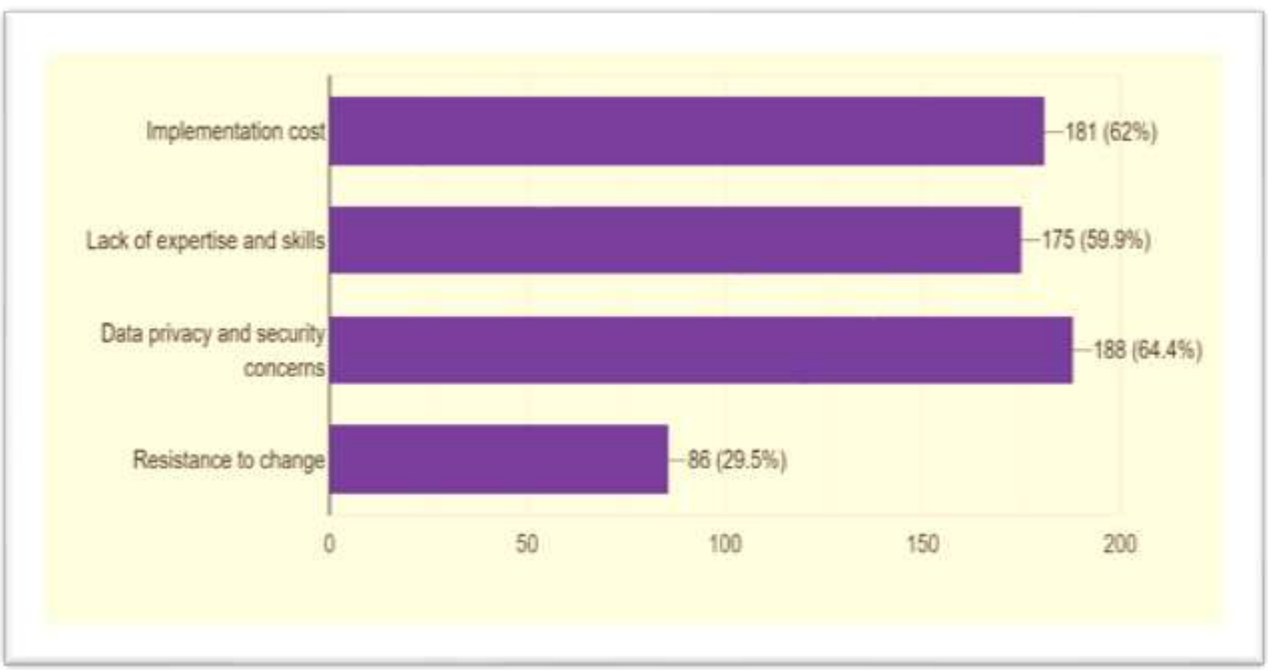
- AI powered data processing and analysis impacted productivity in the organization



➢ AI driven workflow optimization impacted productivity in the organization.



➢ The main challenges the company faced while adopting AI .



Objective 1

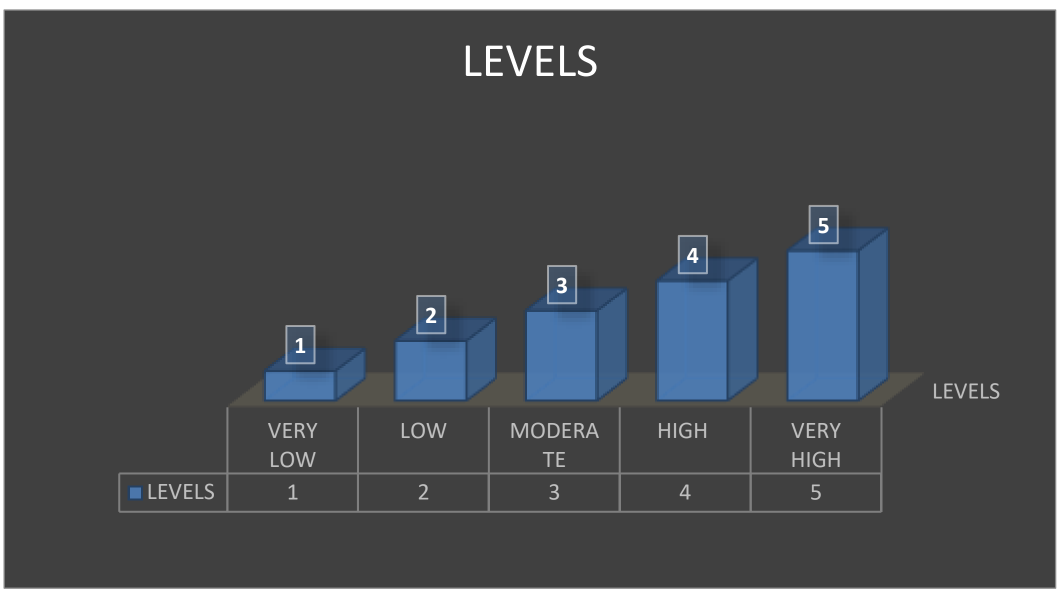
- To study the impact of Artificial Intelligence on the Efficiency of the work in the companies/ corporate offices.

Variables used :-

- ✓ Outcome Variable (Y)– Efficiency of Work

Efficiency of work - The impact of Artificial Intelligence on work efficiency involves leveraging AI technologies to streamline tasks, enhance decision-making, and optimize resource management, leading to increased productivity, reduced operational costs, and improved accuracy in various work processes.

In objective 1 , the **Outcome Variable** as well as the **Predictor Variables** are fitted on **Ordinal Scale** having following levels –
1-Very low
2-Low
3-Moderate
4-High
5-Very high



- ✓ **Predictor variables** –

1. **X1-Workforce** - The impact of Artificial Intelligence on the workforce refers to how AI technologies automate tasks, augment human capabilities, and influence job creation and displacement, thus reshaping skill requirements and employment patterns in various industries.

2. **X2-Capital investment-**Capital investment in artificial intelligence to increase work efficiency involves allocating financial resources to develop and implement AI technologies that automate processes, enhance decision-making, and optimize operations, leading to improved productivity and effectiveness in various business activities.

3. **X3-Employee Productivity-**Employee productivity due to artificial intelligence refers to the enhancement of workers' output and efficiency through the use of AI technologies that automate repetitive tasks, provide data-driven insights, and assist in decision-making, leading to more streamlined and effective work processes.

4. **X4-Workplace Environment**- The impact of Artificial Intelligence on the workplace environment involves the integration of AI technologies to streamline operations, improve productivity, and enhance collaboration among employees, resulting in a more efficient and innovative work environment.

5. **X5-Time Management**- The effect of time management on work efficiency due to Artificial Intelligence refers to the optimization of work schedules, automation of tasks, and prioritization of activities facilitated by AI technologies, leading to increased productivity and effectiveness in completing tasks within specified timeframes.

6. **X6-Workload and Work life Balance**- The effect of workload and work-life balance on work efficiency due to artificial intelligence involves using AI technologies to manage workloads effectively, automate tasks, and promote a better balance between professional responsibilities and personal life, thereby enhancing overall productivity and well-being in the workplace.

7. **X7-Resource Availability** - The impact of Artificial Intelligence on resource availability refers to how AI technologies optimize the allocation and utilization of resources such as manpower, equipment, and materials. By analyzing data and predicting needs, AI helps ensure that resources are allocated efficiently, reducing waste and enhancing productivity.

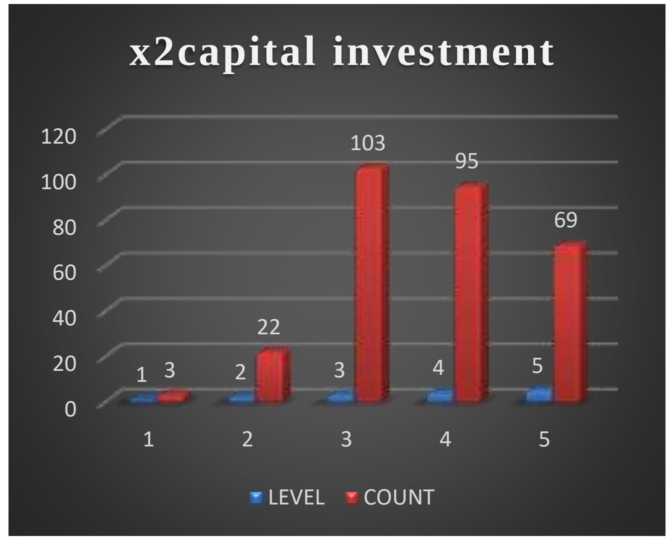
8. **X8-Communication and Collaboration -** The impact of Artificial Intelligence on communication and collaboration involves leveraging AI technologies to facilitate seamless interaction and teamwork among individuals and organizations. AI-powered tools enable improved communication through features such as natural language processing, real-time translation, and voice recognition, while also enhancing collaboration by providing shared workspaces, project management tools, and virtual assistants.

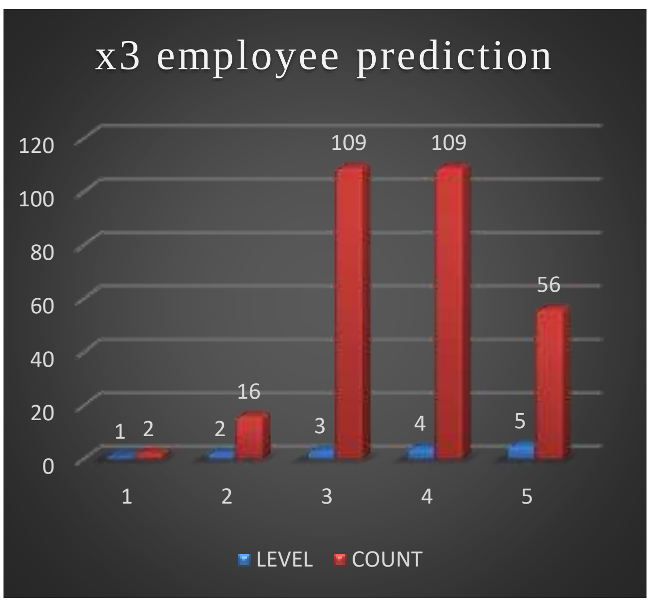
9. **X9-Task Automation**- The impact of Artificial Intelligence on task automation involves using AI technologies to streamline and automate repetitive tasks, freeing up human resources to focus on more complex and strategic activities. AI algorithms analyze data, identify patterns, and make decisions autonomously, enabling the automation of various processes across industries, leading to increased efficiency, reduced errors, and cost savings.

10. **X10-Leadership and Management** - The impact of Artificial Intelligence on leadership and management refers to how AI technologies empower leaders and managers with data-driven insights, predictive analytics, and automation capabilities, enabling them to make informed decisions, optimize resource allocation, foster innovation, and adapt to changing environments more effectively

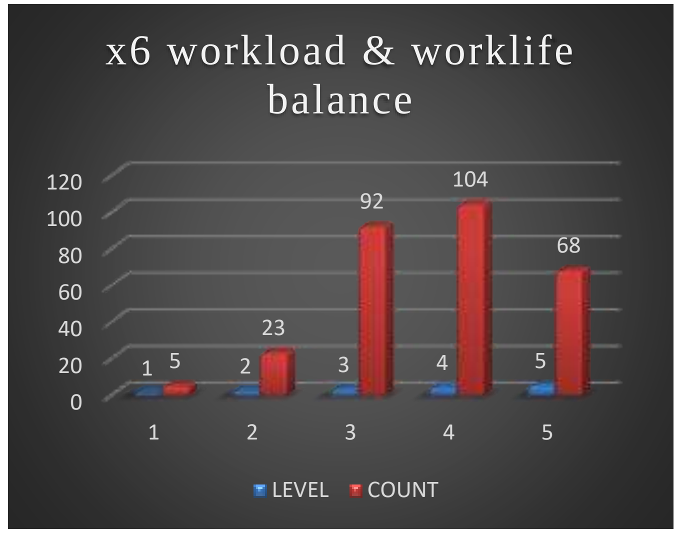
- Graphical representation of impact of AI on predictor

variables described by levels-

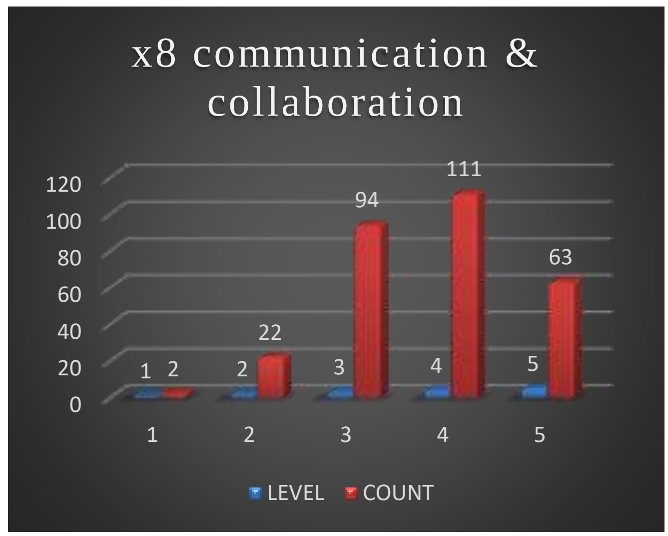


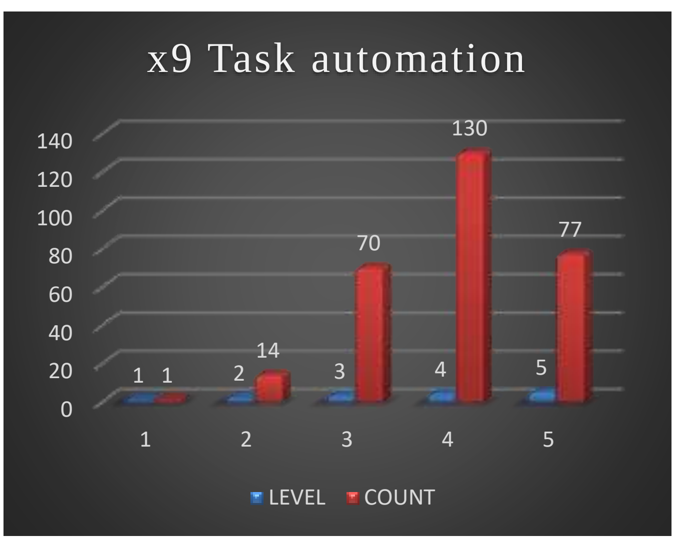



x5 time management
140
120
100
80
60
40
20
0
1 4
2 10
3 79
4 126
5 73
1 2 3 4 5
LEVEL COUNT


x6 workload & worklife balance
120
100
80
60
40
20
0
1 5
2 23
3 92
4 104
5 68
1 2 3 4 5
LEVEL COUNT


x7 resource availability
120
100
80
60
40
20
0
1 2
2 20
3 115
4 99
5 56
1 2 3 4 5
LEVEL COUNT


x8 communication & collaboration
120
100
80
60
40
20
0
1 2
2 22
3 94
4 111
5 63
1 2 3 4 5
LEVEL COUNT


x9 Task automation
140
120
100
80
60
40
20
0
1 1
2 14
3 70
4 130
5 77
1 2 3 4 5
LEVEL COUNT


x10 leadership & management
100
80
60
40
20
0
1 6
2 38
3 90
4 88
5 70
1 2 3 4 5
LEVEL COUNT

Analysis for 1st Objective

- (Full model Fitting)

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Y1 = 1]	-11.762	1.042	127.524	1	<.001	-13.803	-9.720
	[Y1 = 2]	-10.239	.911	126.242	1	<.001	-12.026	-8.453
	[Y1 = 3]	-6.582	.723	82.802	1	<.001	-7.999	-5.164
	[Y1 = 4]	-2.379	.569	17.507	1	<.001	-3.493	-1.264
Location	[X1=1]	-7.478	1.341	31.083	1	<.001	-10.106	-4.849
	[X1=2]	-2.994	.792	14.277	1	<.001	-4.547	-1.441
	[X1=3]	-3.167	.487	42.223	1	<.001	-4.122	-2.212
	[X1=4]	-1.782	.408	19.087	1	<.001	-2.581	-.982
	[X1=5]	0[a]	.	.	0	.	.	.
	[X2=1]	-2.087	1.446	2.083	1	.149	-4.921	.747
	[X2=2]	.245	.659	.138	1	.711	-1.048	1.537
	[X2=3]	-.383	.450	.724	1	.395	-1.264	.499
	[X2=4]	-.266	.407	.427	1	.513	-1.064	.532
	[X2=5]	0[a]	.	.	0	.	.	.
	[X3=1]	2.235	1.778	1.580	1	.209	-1.250	5.721
	[X3=2]	.082	.835	.010	1	.922	-1.554	1.718
	[X3=3]	-.871	.465	3.504	1	.061	-1.782	.041
	[X3=4]	-.207	.426	.235	1	.628	-1.042	.629
	[X3=5]	0[a]	.	.	0	.	.	.
	[X4=1]	.109	2.118	.003	1	.959	-4.041	4.259
	[X4=2]	-1.473	.865	2.902	1	.088	-3.168	.222
	[X4=3]	-.291	.484	.362	1	.547	-1.240	.657
	[X4=4]	-.891	.425	4.396	1	.036	-1.725	-.058
	[X4=5]	0[a]	.	.	0	.	.	.
	[X5=1]	-8.406	1.824	21.241	1	<.001	-11.981	-4.831
	[X5=2]	-2.962	.972	9.280	1	.002	-4.868	-1.056
	[X5=3]	.053	.455	.014	1	.906	-.839	.946
	[X5=4]	-.038	.395	.009	1	.923	-.812	.736
	[X5=5]	0[a]	.	.	0	.	.	.
	[X6=1]	-2.516	1.418	3.150	1	.076	-5.295	.263
	[X6=2]	-.844	.780	1.172	1	.279	-2.374	.685
	[X6=3]	-.410	.473	.752	1	.386	-1.337	.517
	[X6=4]	.432	.384	1.268	1	.260	-.320	1.184
	[X6=5]	0[a]	.	.	0	.	.	.
	[X7=1]	-.865	1.668	.269	1	.604	-4.134	2.404
	[X7=2]	-.322	.757	.180	1	.671	-1.805	1.162
	[X7=3]	.041	.489	.007	1	.934	-.917	.999
	[X7=4]	-.352	.440	.639	1	.424	-1.215	.511

	[X7=5]	0[a]	.	.	0	.	.	.
	[X8=1]	.072	1.883	.001	1	.970	-3.619	3.763
	[X8=2]	.045	.827	.003	1	.957	-1.577	1.666
	[X8=3]	-.715	.476	2.256	1	.133	-1.649	.218
	[X8=4]	-.310	.404	.589	1	.443	-1.101	.481
	[X8=5]	0[a]	.	.	0	.	.	.
	[X9=1]	-18.995	.000	.	1	.	-18.995	-18.995
	[X9=2]	-3.384	.884	14.654	1	<.001	-5.116	-1.651
	[X9=3]	-1.465	.449	10.633	1	.001	-2.345	-.584
	[X9=4]	-.449	.368	1.492	1	.222	-1.170	.272
	[X9=5]	0[a]	.	.	0	.	.	.
	[X10=1]	.414	1.122	.136	1	.712	-1.786	2.614
	[X10=2]	-.068	.637	.011	1	.915	-1.316	1.181
	[X10=3]	-.334	.492	.462	1	.497	-1.298	.630
	[X10=4]	.348	.416	.700	1	.403	-.467	1.163
	[X10=5]	0[a]	.	.	0	.	.	.

Link function: Logit

(Out of the five levels one level is taken as the reference category ,here the level 5 ,i.e. very high is taken as the reference category)

From full model we get 7 variables significant i.e. X1, X2, X3, X4, X5, X6, X9

X1=Workforce

X2=Capital Investment

X3= Employee Prediction

X4=Workplace Environment

X5=Time Management

X6=Workload and Work Life Balance

X9=Task automation

(Before fitting the model equation, we go for the reduced model to check whether the reduced model is as good as full model)

- <u>Proportional odds assumption</u>-

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	421.901			
General	338.753[b]	83.148[c]	120	.996

The above assumption represents parallel line logistic regression model ,i.e.

all Slopes are equal based on cumulative distribution probability of response level. We use score test to test proportional odds assumption, i.e. we test whether or not ordinal restrictions are valid.

We test,

H0: $\beta_{1l} = \beta_1 \quad i.e. \quad \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = \beta_1 \quad \forall \quad l = 1, 2, ., k-1$

H1: $\beta_{1l} \neq \beta_1$ for atleast one l

Where β_1 depicts common slope parameter.

We Hypothesis that there is common slope parameter

<u>Interpretation</u>-

P-value = 0.996 > 0.05

Therefore, we do not reject Null hypothesis and conclude that there is common slope parameter for each outcome variable.

- <u>To test Accuracy</u>

  Confusion matrix-

Y1 * Predicted Response Category Crosstabulation

Count

		Predicted Response Category					Total
		1	2	3	4	5	
Y1	1	3	3	2	0	0	8
	2	1	0	0	0	0	1
	3	4	6	30	9	0	49
	4	0	0	21	110	30	161
	5	0	0	0	20	53	73
Total		8	9	53	139	83	292

Overall Statistics- Accuracy : 0.6712

95% CI : (0.6141, 0.7248)

No Information Rate : 0.476

P-Value [Acc > NIR] : 1.339e-11 , Kappa : 0.4824

Mcnemar's Test P-Value : NA					
Statistics by Class:					
	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	0.37500	0.000000	0.5660	0.7914	0.6386
Specificity	0.98239	0.996466	0.9205	0.6667	0.9043
Pos Pred Value	0.37500	0.000000	0.6122	0.6832	0.7260
Neg Pred Value	0.98239	0.969072	0.9053	0.7786	0.8630
Prevalence	0.02740	0.030822	0.1815	0.4760	0.2842

Detection Rate	0.01027	0.000000	0.1027	0.3767	0.1815
Detection Prevalence	0.02740	0.003425	0.1678	0.5514	0.2500
Balanced Accuracy	0.67870	0.498233	0.7433	0.7290	0.7714

The Accuracy of the Full model is **67.12%**

- (<u>Reduced model fitting</u>)
- PARAMETER ESTIMATES-

Parameter Estimates								
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Y1 = 1]	-11.424	.993	132.455	1	<.001	-13.370	-9.479
	[Y1 = 2]	-9.921	.853	135.356	1	<.001	-11.592	-8.249
	[Y1 = 3]	-6.373	.665	91.751	1	<.001	-7.677	-5.069
	[Y1 = 4]	-2.292	.513	19.942	1	<.001	-3.298	-1.286
Location	[X1=1]	-7.036	1.298	29.394	1	<.001	-9.579	-4.492
	[X1=2]	-3.130	.719	18.926	1	<.001	-4.540	-1.720
	[X1=3]	-3.177	.473	45.113	1	<.001	-4.104	-2.250
	[X1=4]	-1.740	.389	19.992	1	<.001	-2.502	-.977
	[X1=5]	0[a]	.	.	0	.	.	.
	[X4=1]	.111	1.944	.003	1	.954	-3.699	3.922
	[X4=2]	-1.703	.751	5.149	1	.023	-3.174	-.232
	[X4=3]	-.524	.428	1.503	1	.220	-1.362	.314
	[X4=4]	-.812	.388	4.390	1	.036	-1.571	-.052
	[X4=5]	0[a]	.	.	0	.	.	.
	[X5=1]	-7.694	1.541	24.932	1	<.001	-10.715	-4.674
	[X5=2]	-2.691	.916	8.626	1	.003	-4.486	-.895
	[X5=3]	.127	.436	.084	1	.771	-.729	.982
	[X5=4]	-.077	.387	.040	1	.841	-.835	.680
	[X5=5]	0[a]	.	.	0	.	.	.
	[X6=1]	-2.180	1.230	3.144	1	.076	-4.591	.230
	[X6=2]	-.688	.726	.898	1	.343	-2.111	.735
	[X6=3]	-.569	.435	1.710	1	.191	-1.423	.284
	[X6=4]	.366	.370	.979	1	.322	-.359	1.090
	[X6=5]	0[a]	.	.	0	.	.	.
	[X9=1]	-19.344	.000	.	1	.	-19.344	-19.344
	0%[X9=2]	-3.397	.814	17.412	1	<.001	-4.993	-1.802
	[X9=3]	-1.544	.436	12.561	1	<.001	-2.398	-.690
	[X9=4]	-.490	.362	1.830	1	.176	-1.200	.220
	[X9=5]	0[a]	.	.	0	.	.	.
	[X2=1]	-2.084	1.377	2.291	1	.130	-4.782	.614

	[X2=2]	.083	.628	.018	1	.894	-1.147	1.314
	[X2=3]	-.511	.412	1.534	1	.215	-1.319	.297
	[X2=4]	-.400	.380	1.109	1	.292	-1.144	.344
	[X2=5]	0[a]	.	.	0	.	.	.
	[X3=1]	2.176	1.825	1.422	1	.233	-1.400	5.752
	[X3=2]	.088	.811	.012	1	.913	-1.502	1.679
	[X3=3]	-.988	.456	4.702	1	.030	-1.882	-.095
	[X3=4]	-.280	.408	.470	1	.493	-1.080	.520
	[X3=5]	0[a]	.	.	0	.	.	.

Link function: Logit.

- Fitted Ordinal logistic model-

$$\Pi_1 = \frac{e^{logit\ F1}}{1 + e^{logitF1}}$$

where,

LogitF1=-11.424-7.036*Workforce(1)-3.130*Workforce(2) -3.177*Workforce(3)-1.740*Workforce(4) -2.084*Capital investment(1)+0.083*Capital investment(2)-0.511*Capital investment(3)-0.400*Capital investment(4)+2.176*Employee productivity(1)+0.088* Employee productivity(2)-0.988* Employee productivity(3)-0.280* Employee productivity(4)+0.111*Workplace environment(1)-1.703*Workplace environment(2)-0.524*Workplace environment(3)-0.812*Workplace environment(4)-7.694*Time management(1)-2.691*Time management(2)0.127*Time management(3)-0.077*Time management(4)-2.180*Workload and Work life balance(1)-0.688*Workload and Work life balance(2)-0.569*Workload and Work life balance(3)+0.366*Workload and Work life balance(4)-19.344*Task automation(1)-3.397*Task automation(2)-1.544*Task automation(3)-4.90*Task automation(4)

$$\Pi_2 = \frac{e^{logit\ F2}}{1 + e^{logitF2}}$$

where,

LogitF2 =-9.921-7.036*Workforce(1)-3.130*Workforce(2) -3.177*Workforce(3)-1.740*Workforce(4) -2.084*Capital investment(1)+0.083*Capital investment(2)-0.511*Capital investment(3)-0.400*Capital investment(4)+2.176*Employee productivity(1)+0.088* Employee productivity(2)-0.988* Employee productivity(3)-0.280* Employee productivity(4)+0.111*Workplace environment(1)-1.703*Workplace environment(2)-0.524*Workplace environment(3)-0.812*Workplace environment(4)-7.694*Time management(1)-2.691*Time management(2)0.127*Time management(3)-0.077*Time management(4)-2.180*Workload and Work life balance(1)-0.688*Workload and Work life balance(2)-0.569*Workload and Work life balance(3)+0.366*Workload and Work life balance(4)-19.344*Task automation(1)-3.397*Task automation(2)-1.544*Task automation(3)-4.90*Task automation(4)

$$\Pi_3 = \frac{e^{logit\ F3}}{1 + e^{logitF3}}$$

where,

LogitF3 =-6.373-7.036*Workforce(1)-3.130*Workforce(2) -3.177*Workforce(3)-1.740*Workforce(4) -2.084*Capital investment(1)+0.083*Capital investment(2)-0.511*Capital investment(3)-0.400*Capital investment(4)+2.176*Employee productivity(1)+0.088* Employee productivity(2)-0.988* Employee productivity(3)-0.280* Employee productivity(4)+0.111*Workplace environment(1)-1.703*Workplace environment(2)-0.524*Workplace environment(3)-0.812*Workplace environment(4)-7.694*Time management(1)-2.691*Time management(2)0.127*Time management(3)-0.077*Time management(4)-2.180*Workload and Work life balance(1)-0.688*Workload and Work life balance(2)-0.569*Workload and Work life balance(3)+0.366*Workload and Work life balance(4)-19.344*Task automation(1)-3.397*Task automation(2)-1.544*Task automation(3)-4.90*Task automation(4)

$$\Pi_4 = \frac{e^{logit\ F4}}{1 + e^{logitF4}}$$

where,

LogitF4 =-2.292-7.036*Workforce(1)-3.130*Workforce(2) -3.177*Workforce(3)-1.740*Workforce(4) -2.084*Capital investment(1)+0.083*Capital investment(2)-0.511*Capital investment(3)-0.400*Capital investment(4)+2.176*Employee productivity(1)+0.088* Employee productivity(2)-0.988* Employee productivity(3)-0.280* Employee productivity(4)+0.111*Workplace environment(1)-1.703*Workplace environment(2)-0.524*Workplace environment(3)-0.812*Workplace environment(4)-7.694*Time management(1)-2.691*Time management(2)0.127*Time management(3)-0.077*Time management(4)-2.180*Workload and Work life balance(1)-0.688*Workload and Work life balance(2)-0.569*Workload and Work life balance(3)+0.366*Workload and Work life balance(4)-19.344*Task automation(1)-3.397*Task automation(2)-1.544*Task automation(3)-4.90*Task automation(4)

- Proportional odds assumption-

Test of Parallel Lines[a]				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	417.499			
General	351.845[b]	65.654[c]	84	.931

We test,

H0: $\beta_{1l} = \beta_1 \;\; i.e. \;\; \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = \beta_1 \quad \forall \quad l = 1, 2, \ldots, k-1$

H1: $\beta_{1l} \neq \beta_1$ for atleast one l

Where β_1 depicts common slope parameter.

We Hypothesis that there is common slope parameter

Interpretation-P-value = 0.931 > 0.05

Therefore, we do not reject Null hypothesis and conclude that there is common slope parameter for each outcome variable.

- Model fitting information and Goodness-of-fit test-

Model Fitting Information				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	699.945			
Final	417.499	282.446	28	<.001
Link function: Logit.				

This table provides information on the **-2 Log Likelihood** of the null model (without predictors) and the final model (with predictors), with lower values suggesting a better fit. Additionally, the chi-square statistic and associated p-value assess whether the model significantly improves fit compared to a null model, with a p-value less than 0.05 indicating a significant improvement.

Hypothesis:
H_0: The predictor variables do not significantly improve the fit of the model compared to the null model (without predictors)
H_1: At least one of the predictor variables significantly improves the fit of the model
Interpretation: P-value = 0.00 < 0.05 .Therefore, we reject Null hypothesis and say that the predictor variables significantly improves the fit of the model.

- Goodness of fit test-

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	680.588	884	1.000
Deviance	404.288	884	1.000
Link function: Logit.			

This includes tests like Pearson and Deviance, which assess how well the model fits the data. If p-value > 0.05 it indicates a good fit.
Hypothesis:H_0: The model fits the data well.
H_1: The model does not fit data well

Interpretation –Here P-value – 1.000 >0.05.Therefore, we do not reject Null hypothesis and say that the model fits the data well.

➢ To test accuracy

Confusion Matrix-

Y1 * Predicted Response Category Crosstabulation							
Count							
		Predicted Response Category					Total
		1	2	3	4	5	
Y1	1	2	3	2	0	0	7
	2	2	0	0	0	0	2
	3	4	6	30	3	0	43
	4	0	0	19	117	31	167
	5	0	0	2	19	52	73
Total		8	9	53	139	83	292

Overall Statistics-Accuracy : 0.6884 ,95% CI : (0.6318, 0.741),No Information Rate : 0.476,P-Value [Acc > NIR]-1.842e-13 ,Kappa :0.5046,Mcnemar's Test P-Value : NA

Statistics by Class:					
	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	**0.250000**	**0.000000**	**0.5660**	**0.8417**	**0.6265**
Specificity	**0.982394**	**0.992933**	**0.9456**	**0.6732**	**0.8995**
Pos Pred Value	**0.285714**	**0.000000**	**0.6977**	**0.7006**	**0.7123**
Neg Pred Value	**0.978947**	**0.968966**	**0.9076**	**0.8240**	**0.8584**
Prevalence	**0.027397**	**0.030822**	**0.1815**	**0.4760**	**0.2842**
Detection Rate	**0.006849**	**0.000000**	**0.1027**	**0.4007**	**0.1781**
Detection Prevalence	**0.023973**	**0.006849**	**0.1473**	**0.5719**	**0.2500**
Balanced Accuracy	**0.616197**	**0.496466**	**0.7558**	**0.7575**	**0.7630**

The accuracy of the Reduced model is **68.84 %**

❖ To check Multicollinearity-

	GVIF	Df	GVIF^(1/(2*Df))
Y1	**4.241951**	**4**	**1.197969**
X1	**1.867255**	**1**	**1.366475**

X2	**2.799336**	**4**	**1.137318**
X3	**3.012751**	**4**	**1.147811**
X4	**2.437153**	**1**	**1.561138**
X5	**1.877715**	**1**	**1.370297**
X6	**2.368867**	**1**	**1.539112**
X7	**3.813546**	**4**	**1.182132**
X8	**3.871013**	**4**	**1.184345**
X9	**1.786982**	**1**	**1.336781**
X10	**3.817951**	**4**	**1.182303**

GVIF (Generalized Variance Inflation Factor) is a measure used to detect multicollinearity in regression models, particularly when dealing with categorical predictors. It extends the concept of **VIF** (Variance Inflation Factor) to handle the complexity of models with categorical variables.GVIF can be adjusted for the number of degrees of freedom (df) associated with the predictor, resulting in a measure called **GVIF^(1/(2*df))**. This adjustment helps interpret GVIF in models with categorical variables having multiple levels.
GVIF^(1/(2*df)) close to 1: Indicates little to no multicollinearity.
GVIF^(1/(2*df)) greater than 1: Indicates the presence of

multicollinearity, with higher values suggesting more severe multicollinearity.

Interpretation:
In the above table all the adjusted GVIF values are close to 1 which indicates relatively little to no collinearity which is acceptable.

- Likelihood Ratio test-

Model 1: Y1 ~ X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 + X9 + X10
Model 2: Y1 ~ X1 + X2 + X3 + X4 + X5 + X6 + X9

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	**44**	**-212.34**			
2	**32**	**-216.90**	**12**	**9.1206**	**0.6926**

Hypothesis :H_0:Reduced model is as good as full model
H_1 :Reduced model is not as good as full model

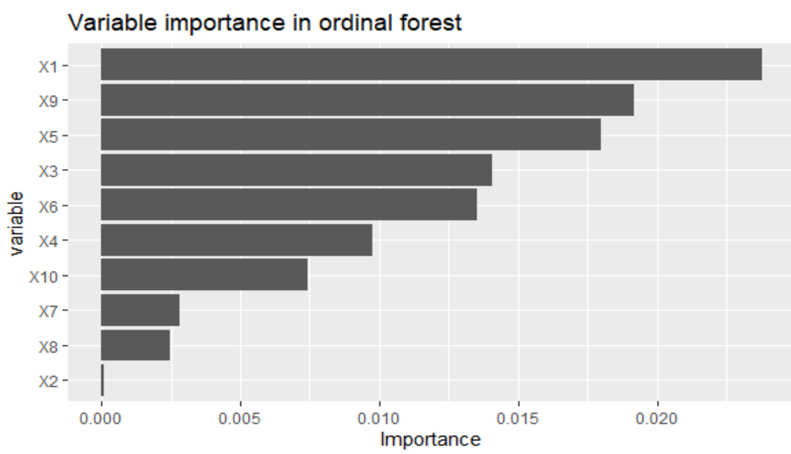
Interpretation: Here the P-value =0.6926>0.05.
Hence, we do not reject the null hypothesis and conclude that Reduced model is as good as Full model.

- Analysis done Ordinal Forest technique

	Length	Class	Mode
forestfinal	13	ranger	list
bordersbest	6	-none-	numeric
forests	100	-none-	list
perffunctionvalues	100	-none-	numeric
bordersb	600	-none-	numeric
classes	5	-none-	character
nsets	1	-none-	numeric
ntreeperdiv	1	-none-	numeric
ntreefinal	1	-none-	numeric
perffunction	1	-none-	character
classimp	1	-none-	logical
nbest	1	-none-	numeric
classfreq	5	table	numeric
varimp	10	-none-	numeric

Classes of ordinal target variable:
"1", "2", "3", "4", "5"
Importance of individual variable in the ordinal forest model

	variable	Importance
X1	X1	2.378499e-02
X2	X2	9.950023e-05
X3	X3	1.405385e-02
X4	X4	9.771447e-03
X5	X5	1.796626e-02
X6	X6	1.353314e-02
X7	X7	2.828807e-03
X8	X8	2.458726e-03
X9	X9	1.916054e-02
X10	X10	7.419457e-03



Interpretation-From the table of importance and graph we can see that the variable X1,X9,X5 has the most importance compared to rest of the variables.

- Training accuracy of the above fitted model
  Confusion matrix

pred_labels	1	2	3	4	5
1	3	0	0	0	0
2	3	6	0	0	0
3	0	0	39	3	0
4	0	0	1	84	0
5	0	0	0	11	62
Accuracy=0.9150943					

Accuracy of the training model is 91.50%

- Testing accuracy of the above fitted model

Confusion matrix

pred_labels	1	2	3	4	5
1	0	0	0	0	0
2	0	2	0	0	0
3	2	0	7	2	0
4	0	1	5	26	2
5	0	0	1	13	19
Accuracy=0.675					

Accuracy of the testing model is 67.50%

Objective 2

✓ To analyze the impact of Artificial Intelligence on the Expenditure of the various companies.

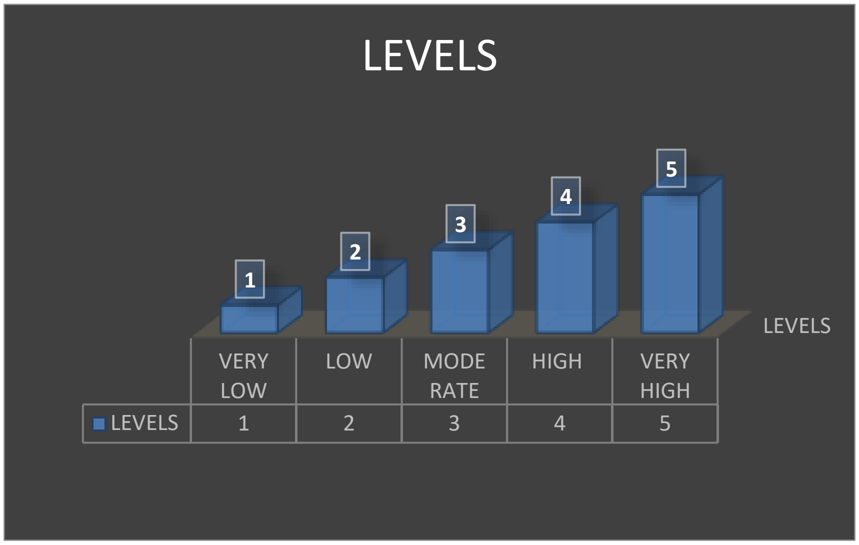
Variables used :-

- ✓ Outcome Variable (Y)- Expenditure

Expenditure- Artificial intelligence can have a significant impact on expenditure by streamlining processes, reducing operational costs, and optimizing resource allocation. Impact of AI on expenditure depends on how effectively it is implemented and integrated into various aspects of the company's operations.

In the objective 2 the **Outcome Variable** as well as the **Predictor Variables** are fitted on **Ordinal Scale** having following levels –

1-Very low
2-Low
3-Moderate
4-High
5-Very high



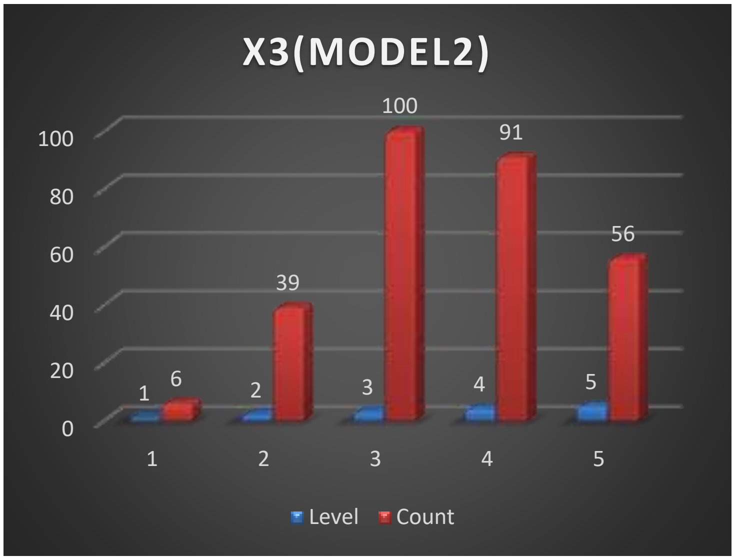
✓ **<u>Predictor variables</u> –**

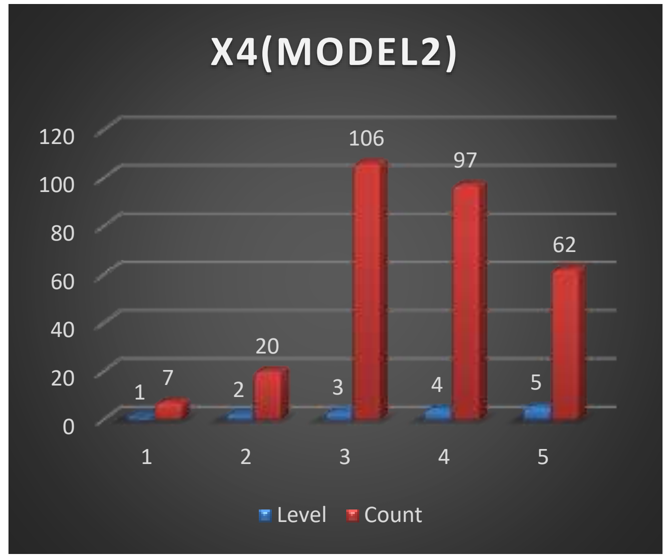
1. <u>XI- Resource Management</u> - AI has a significant impact on resource management by optimizing allocation, predicting demand, and improving efficiency across various sectors like energy, water, agriculture, and transportation.
2. <u>X2- Strategic sourcing-</u>AI has significantly impacted strategic sourcing by enhancing efficiency, accuracy, and decision-making processes. AI also automates repetitive tasks, freeing up procurement professionals to focus on strategic activities.
3. <u>X3- Labour cost Management</u> - AI can have a profound impact on labor cost management by automating repetitive tasks, optimizing workforce scheduling, and identifying areas for efficiency improvements.
4.<u>X4-Risk Management</u>- AI revolutionizes risk management by enhancing predictive analytics, automating processes, and improving decision-

making.AI-powered systems can automate routine risk assessment tasks, such as credit scoring or fraud detection, reducing human error and increasing efficiency.
5. <u>X5- Customer Relationship Management</u> - AI has revolutionized customer relationship management (CRM) by enabling businesses to analyze vast amounts of customer data, personalize interactions, and automate processes.AI enhances CRM by fostering deeper customer engagement and loyalty while streamlining operations for businesses.
6. <u>X6-Supply Chain Management</u> - AI can optimize supply chain operations, reducing inventory holding costs and minimizing stockouts, leading to more efficient expenditure on inventory management.

7. X7-Operational Efficiency - AI can automate tasks, reducing the need for manual labor and thus lowering operational costs.
8. X8-Energy Efficiency - AI algorithms can analyze real-time data from sensors and meters to optimize energy consumption in buildings, factories, and transportation systems. This includes adjusting lighting, heating, cooling, and other systems based on occupancy, weather conditions, and energy prices to minimize waste and reduce costs.
9. X9-Product Design - AI has revolutionized product design by enabling faster prototyping, enhanced customization, and predictive analytics to anticipate user preferences. It streamlines the design process, optimizes product performance, and fosters innovation through data-driven insights.

Graphical Representation of the impact of AI on predictor variables described by levels-


X3(MODEL2)
100
80
60
40
20
0
100
91
56
39
1 6
2
3
4
5
1
2
3
4
5
Level Count

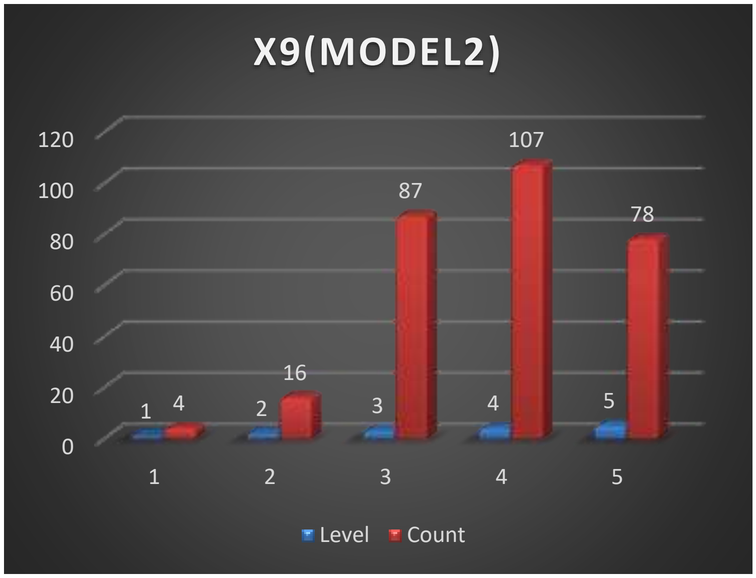

X4(MODEL2)
120
100
80
60
40
20
0
106
97
62
20
1 7
2
3
4
5
1
2
3
4
5
Level Count


X5(MODEL2)
120
100
80
60
40
20
0
104
89
70
27
1 2
2
3
4
5
1
2
3
4
5
Level Count


X6(MODEL2)
120
100
80
60
40
20
0
113
89
75
1 5
2 10
3
4
5
1
2
3
4
5
Level Count


X7(MODEL2)
120
100
80
60
40
20
0
112
105
57
12
1 6
2
3
4
5
1
2
3
4
5
Level Count


X8(MODEL2)
120
100
80
60
40
20
0
113
88
75
11
1 5
2
3
4
5
Level Count
1
2
3
4
5



Analysis of 2nd objective

(To increase the accuracy of the model we divide data into 80% and 20%)

- (Full model fitting)-
- For training data:

Parameter Estimates								
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Y2 = 1]	-10.473	1.049	99.743	1	<.001	-12.528	-8.418
	[Y2 = 2]	-8.142	.834	95.372	1	<.001	-9.776	-6.508
	[Y2 = 3]	-4.574	.719	40.425	1	<.001	-5.984	-3.164
	[Y2 = 4]	-2.201	.649	11.508	1	<.001	-3.473	-.929
Location	[X1=1]	-10.831	3.182	11.584	1	<.001	-17.068	-4.594
	[X1=2]	-2.146	.811	6.996	1	.008	-3.736	-.556
	[X1=3]	-1.189	.590	4.064	1	.044	-2.345	-.033
	[X1=4]	-.756	.527	2.060	1	.151	-1.788	.276
	[X1=5]	0^a	.	.	0	.	.	.
	[X2=1]	8.250	3.050	7.315	1	.007	2.271	14.228
	[X2=2]	1.180	.817	2.086	1	.149	-.421	2.781
	[X2=3]	.510	.518	.971	1	.324	-.505	1.525
	[X2=4]	.412	.444	.864	1	.353	-.457	1.282
	[X2=5]	0^a	.	.	0	.	.	.
	[X3=1]	-6.152	1.563	15.500	1	<.001	-9.214	-3.089
	[X3=2]	-1.318	.684	3.711	1	.054	-2.658	.023
	[X3=3]	-.590	.519	1.294	1	.255	-1.607	.427

	[X3=4]	.310	.461	.453	1	.501	-.594	1.215
	[X3=5]	0[a]	.	.	0	.	.	.
	[X4=1]	1.724	1.121	2.365	1	.124	-.473	3.921
	[X4=2]	-.759	.785	.933	1	.334	-2.297	.780
	[X4=3]	-.775	.488	2.522	1	.112	-1.732	.181
	[X4=4]	.097	.466	.043	1	.835	-.817	1.011
	[X4=5]	0[a]	.	.	0	.	.	.
	[X5=1]	0[a]	.	.	0	.	.	.
	[X5=2]	-2.401	.757	10.048	1	.002	-3.885	-.916
	[X5=3]	-1.091	.534	4.172	1	.041	-2.138	-.044
	[X5=4]	-.606	.464	1.711	1	.191	-1.515	.302
	[X5=5]	0[a]	.	.	0	.	.	.
	[X6=1]	.363	1.594	.052	1	.820	-2.761	3.487
	[X6=2]	-3.598	.973	13.672	1	<.001	-5.505	-1.691
	[X6=3]	-1.007	.499	4.075	1	.044	-1.984	-.029
	[X6=4]	-1.224	.438	7.804	1	.005	-2.082	-.365
	[X6=5]	0[a]	.	.	0	.	.	.
	[X7=1]	1.257	1.367	.845	1	.358	-1.423	3.936
	[X7=2]	1.160	1.078	1.158	1	.282	-.952	3.272
	[X7=3]	.715	.528	1.833	1	.176	-.320	1.750
	[X7=4]	-.047	.474	.010	1	.921	-.976	.882
	[X7=5]	0[a]	.	.	0	.	.	.
	[X8=1]	-2.979	1.600	3.465	1	.063	-6.115	.158
	[X8=2]	-1.992	1.193	2.786	1	.095	-4.330	.347
	[X8=3]	-2.376	.499	22.697	1	<.001	-3.354	-1.399
	[X8=4]	-1.216	.435	7.805	1	.005	-2.069	-.363
	[X8=5]	0[a]	.	.	0	.	.	.
	[X9=1]	-4.328	1.473	8.628	1	.003	-7.216	-1.440
	[X9=2]	1.753	.830	4.459	1	.035	.126	3.380
	[X9=3]	-.030	.536	.003	1	.955	-1.080	1.020
	[X9=4]	.197	.411	.231	1	.631	-.607	1.002
	[X9=5]	0[a]	.	.	0	.	.	.

- ❖ Training accuracy

Confusion Matrix

Y2 * Predicted Response Category Crosstabulation							
Count							
		Predicted Response Category					Total
		1	2	3	4	5	
Y2	1	3	1	2	0	0	6
	2	0	6	6	1	0	13
	3	0	2	55	11	6	74
	4	0	0	20	35	13	68
	5	0	0	1	16	56	73
Total		3	9	84	63	75	234

Overall Statistics- Accuracy : 0.6624 ,95% CI : (0.5979, 0.7227) ,No Information Rate : 0.3162 ,P-Value [Acc > NIR] : < 2.2e-16 ,Kappa : 0.5217,Mcnemar's Test P-Value : NA

Statistics by Class:					
	Class: 1	**Class: 2**	**Class: 3**	**Class: 4**	**Class: 5**
Sensitivity	**0.50000**	**0.46154**	**0.7432**	**0.5147**	**0.7671**
Specificity	**1.00000**	**0.98643**	**0.8187**	**0.8313**	**0.8820**
Pos Pred Value	**1.00000**	**0.66667**	**0.6548**	**0.5556**	**0.7467**
Neg Pred Value	**0.98701**	**0.96889**	**0.8733**	**0.8070**	**0.8931**
Prevalence	**0.02564**	**0.05556**	**0.3162**	**0.2906**	**0.3120**
Detection Rate	**0.01282**	**0.02564**	**0.2350**	**0.1496**	**0.2393**
Detection Prevalence	**0.01282**	**0.03846**	**0.3590**	**0.2692**	**0.3205**
Balanced Accuracy	**0.75000**	**0.72398**	**0.7810**	**0.6730**	**0.8246**
The accuracy of the training model is **66.24%**					

- For testing data:

Confusion Matrix

Y2 * Predicted Response Category Crosstabulation							
Count							
		Predicted Response Category					Total
		1	2	3	4	5	
Y2	1	0	0	0	0	0	0
	2	0	1	0	0	0	1
	3	0	0	5	7	0	12
	4	0	0	7	8	6	21
	5	0	0	2	4	18	24
Total		0	1	14	19	24	58

Overall Statistics-Accuracy : 0.5517,95% CI : (0.4154, 0.6826),No Information Rate : 0.4138,P-Value [Acc > NIR] : 0.02354,Kappa : 0.3207Mcnemar's Test P-Value : NA

Statistics by Class:					
	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	NA	1.00000	0.35714	0.4211	0.7500
Specificity	1	1.00000	0.84091	0.6667	0.8235
Pos Pred Value	NA	1.00000	0.41667	0.3810	0.7500
Neg Pred Value	NA	1.00000	0.80435	0.7027	0.8235
Prevalence	0	0.01724	0.24138	0.3276	0.4138
Detection Rate	0	0.01724	0.08621	0.1379	0.3103
Detection Prevalence	0	0.01724	0.20690	0.3621	0.4138
Balanced Accuracy	NA	1.00000	0.59903	0.5439	0.7868

Testing accuracy of the testing data=**55.17%**

- Checking proportional odds assumption of full model:

Hypothesis-H_0: $\beta_{1l} = \beta_1 \quad \forall \quad l = 1, 2, 3, 4$

H_1: $\beta_{1l} \neq \beta_1$ for atleast one l

We Hypothesis that there is common slope parameter for each of the cumulative logit regression equation instead of 4 distinct slopes.

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	384.121			
General	297.635[b]	86.486[c]	105	.906

Interpretation-P-value = 0.906 > 0.05.Therefore, we do not reject Null hypothesis and conclude that there is common slope parameter for each outcome variable.

❖ (Reduced model fitting):

(For reduced model we only consider significant variables from the full model)

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Y2 = 1]	-10.281	1.018	101.892	1	<.001	-12.277	-8.285
	[Y2 = 2]	-7.890	.781	102.158	1	<.001	-9.420	-6.360
	[Y2 = 3]	-4.483	.668	44.982	1	<.001	-5.793	-3.173
	[Y2 = 4]	-2.235	.600	13.898	1	<.001	-3.410	-1.060
Location	[X1=1]	-10.300	3.148	10.705	1	.001	-16.470	-4.130
	[X1=2]	-2.179	.768	8.047	1	.005	-3.685	-.674
	[X1=3]	-1.326	.559	5.631	1	.018	-2.421	-.231
	[X1=4]	-.793	.498	2.536	1	.111	-1.769	.183
	[X1=5]	0[a]	.	.	0	.	.	.
	[X2=1]	7.667	2.882	7.080	1	.008	2.020	13.315
	[X2=2]	.710	.774	.842	1	.359	-.807	2.227
	[X2=3]	.451	.503	.802	1	.370	-.536	1.437
	[X2=4]	.279	.427	.427	1	.513	-.558	1.115
	[X2=5]	0[a]	.	.	0	.	.	.
	[X3=1]	-5.413	1.466	13.624	1	<.001	-8.287	-2.539
	[X3=2]	-1.365	.634	4.635	1	.031	-2.609	-.122
	[X3=3]	-.462	.504	.840	1	.359	-1.450	.526
	[X3=4]	.231	.445	.269	1	.604	-.642	1.104
	[X3=5]	0[a]	.	.	0	.	.	.
	[X5=1]	0[a]	.	.	0	.	.	.
	[X5=2]	-1.835	.665	7.606	1	.006	-3.139	-.531
	[X5=3]	-.969	.509	3.630	1	.057	-1.966	.028
	[X5=4]	-.497	.446	1.243	1	.265	-1.372	.377
	[X5=5]	0[a]	.	.	0	.	.	.
	[X6=1]	.050	1.534	.001	1	.974	-2.956	3.057
	[X6=2]	-3.623	.901	16.173	1	<.001	-5.389	-1.857
	[X6=3]	-1.000	.472	4.489	1	.034	-1.924	-.075
	[X6=4]	-1.103	.429	6.601	1	.010	-1.945	-.262
	[X6=5]	0[a]	.	.	0	.	.	.
	[X8=1]	-2.391	1.382	2.995	1	.083	-5.099	.317
	[X8=2]	-1.796	1.138	2.489	1	.115	-4.027	.435
	[X8=3]	-2.441	.482	25.694	1	<.001	-3.385	-1.497
	[X8=4]	-1.183	.408	8.396	1	.004	-1.984	-.383
	[X8=5]	0[a]	.	.	0	.	.	.
	[X9=1]	-3.328	1.389	5.737	1	.017	-6.051	-.605
	[X9=2]	1.668	.780	4.576	1	.032	.140	3.196
	[X9=3]	.127	.511	.062	1	.804	-.874	1.128
	[X9=4]	.145	.401	.131	1	.718	-.640	.930
	[X9=5]	0[a]	.	.	0	.	.	.

Ordinal logistic regression model based on parameter estimates by the reduced model.

$$\Pi_1 = \frac{e^{logit\,F1}}{1 + e^{logitF1}}$$

Where , **logit F1**= -10.281-10.300*Resource management1- 2.179* Resource management 2- 1.326*Resource management3-0.793* Resource management4+7.667*Strategic sourcing1+0.710 *Strategic sourcing2+0.451 *Strategic sourcing3+0.279 *Strategic sourcing4-5.413*Labor cost management1-1.365* Labor cost management2-0.462* Labor cost management3+0.231*Labor cost management4-1.835*Customer relationship management2-0.969* Customer relationship management3-0.497 *Customer relationship management4+0.50*Supply chain management1-3.623* Supply chain management2-1.000* Supply chain management3-1.103 *Supply chain management4-2.391*Energy efficiency1-1.796* Energy efficiency2-2.441* Energy efficiency3-1.183 *Energy efficiency4-3.328* Product design1+1.668* Product design2+0.127* Product design3+0.145 *Product design4

$$\Pi_2 = \frac{e^{logit\,F2}}{1 + e^{logitF2}}$$

Where , **logit F2**= -7.890-10.300*Resource management1- 2.179* Resource management 2- 1.326*Resource management3-0.793* Resource management4+7.667*Strategic sourcing1+0.710 *Strategic sourcing2+0.451 *Strategic sourcing3+0.279 *Strategic sourcing4-5.413*Labor cost management1-1.365* Labor cost management2-0.462* Labor cost management3+0.231*Labor cost management4-1.835*Customer relationship management2-0.969* Customer relationship management3-0.497 *Customer relationship management4+0.50*Supply chain management1-3.623* Supply chain management2-1.000* Supply chain management3-1.103 *Supply chain management4-2.391*Energy efficiency1-1.796* Energy efficiency2-2.441* Energy efficiency3-1.183 *Energy efficiency4-3.328* Product design1+1.668* Product design2+0.127* Product design3+0.145 *Product design4

$$\Pi_3 = \frac{e^{logit\ F3}}{1 + e^{logitF3}}$$

Where , **logit F3**= -4.483-10.300*Resource management1- 2.179* Resource management 2- 1.326*Resource management3-0.793* Resource management4+7.667*Strategic sourcing1+0.710 *Strategic sourcing2+0.451 *Strategic sourcing3+0.279 *Strategic sourcing4-5.413*Labor cost management1-1.365* Labor cost management2-0.462* Labor cost management3+0.231*Labor cost management4-1.835*Customer relationship management2-0.969* Customer relationship management3-0.497 *Customer relationship management4+0.50*Supply chain management1-3.623* Supply chain management2-1.000* Supply chain management3-1.103 *Supply chain management4-2.391*Energy efficiency1-1.796* Energy efficiency2-2.441* Energy efficiency3-1.183 *Energy efficiency4-3.328* Product design1+1.668* Product design2+0.127* Product design3+0.145 *Product design4

$$\Pi_4 = \frac{e^{logit\ F4}}{1 + e^{logitF4}}$$

Where , **logit F4**= -2.235-10.300*Resource management1- 2.179* Resource management 2- 1.326*Resource management3-0.793* Resource management4+7.667*Strategic sourcing1+0.710 *Strategic sourcing2+0.451 *Strategic sourcing3+0.279 *Strategic sourcing4-5.413*Labor cost management1-1.365* Labor cost management2-0.462* Labor cost management3+0.231*Labor cost management4-1.835*Customer relationship management2-0.969* Customer relationship management3-0.497 *Customer relationship management4+0.50*Supply chain management1-3.623* Supply chain management2-1.000* Supply chain management3-1.103 *Supply chain management4-2.391*Energy efficiency1-1.796* Energy efficiency2-2.441* Energy efficiency3-1.183 *Energy efficiency4-3.328* Product design1+1.668* Product design2+0.127* Product design3+0.145 *Product design4

- Training accuracy:

Confusion matrix

Y2 * Predicted Response Category Crosstabulation							
Count							
		Predicted Response Category					Total
		1	2	3	4	5	
Y2	1	3	0	0	0	0	3
	2	1	6	2	0	0	9
	3	2	6	55	20	1	84
	4	0	1	11	35	16	63
	5	0	0	6	13	56	75
Total		6	13	74	68	73	234

Overall Statistics- Accuracy : 0.6624 ,95% CI : (0.5979, 0.7227), No Information Rate : 0.3162, P-Value [Acc > NIR] : < 2.2e-16 ,Kappa : 0.5217,Mcnemar's Test P-Value : NA

Statistics by Class:					
	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	0.5 0000	0.46154	0.7432	0.5147	0.7671
Specificity	1.00000	0.98643	0.8187	0.8313	0.8820
Pos Pred Value	1.00000	0.66667	0.6548	0.5556	0.7467
Neg Pred Value	0.98701	0.96889	0.8733	0.8070	0.8931
Prevalence	0.02564	0.05556	0.3162	0.2906	0.3120
Detection Rate	0.01282	0.02564	0.2350	0.1496	0.2393
Detection Prevalence	0.01282	0.03846	0.3590	0.2692	0.3205
Balanced Accuracy	0.75000	0.72398	0.7810	0.6730	0.8246

Overall training accuracy of the reduced model is **66.24%**

- For testing data: Confusion matrix

Y2 * Predicted Response Category Crosstabulation							
Count							
		Predicted Response Category					Total
		1	2	3	4	5	
Y2	1	0	0	0	0	0	0
	2	0	1	0	0	0	1
	3	0	0	5	7	0	12
	4	0	0	7	8	6	21
	5	0	0	2	4	18	24
Total		0	1	14	19	24	58

Overall Statistics-Accuracy : 0.5517 ,95% CI : (0.4154, 0.6826), No Information Rate : 0.4138 P-Value [Acc > NIR] : 0.02354 ,Kappa : 0.3207,Mcnemar's Test P-Value : NA

Statistics by Class:					
	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	NA	1.00000	0.35714	0.4211	0.7500
Specificity	1	1.00000	0.84091	0.6667	0.8235
Pos Pred Value	NA	1.00000	0.41667	0.3810	0.7500
Neg Pred Value	NA	1.00000	0.80435	0.7027	0.8235
Prevalence	0	0.01724	0.24138	0.3276	0.4138
Detection Rate	0	0.01724	0.08621	0.1379	0.3103
Detection Prevalence	0	0.01724	0.20690	0.3621	0.4138
Balanced Accuracy	NA	1.00000	0.59903	0.5439	0.7868

Overall testing accuracy of the reduced model is **55.17%**

- Checking proportional odds assumption of reduced model:

  Hypothesis-H_0: $\beta_{1l} = \beta_1 \quad \forall \quad l = 1, 2, 3, 4$

  H_1: $\beta_{1l} \neq \beta_1$ for atleast one l

  We Hypothesis that there is common slope parameter for each of the cumulative logit regression equation instead of 4 distinct slopes.

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	382.511			
General	302.398[b]	80.113[c]	81	.507

Interpretation-P-value = 0.507 > 0.05.Therefore, we do not reject Null hypothesis and conclude that there is common slope parameter for each outcome variable.

- Model Fitting Information

  Hypothesis:

  H_0: The predictor variables do not significantly improve the fit of the model compared to the null model (without predictors)

  H_1: At least one of the predictor variables significantly improves the fit of the model

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
InterceptOnly	590.377			
Final	382.511	207.866	27	<.001

Interpretation:P-value = 0.001 < 0.05.Therefore, we reject Null hypothesis and say the predictor variables significantly improves the fit of the model.

- <u>Goodness-of-Fit Tests</u>**:**
  Hypothesis:H_0: The model fits the data well.
  H_1: The model does not fit data well

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	513.110	665	1.000
Deviance	359.465	665	1.000

<u>Interpretation</u>-P-value = 1.0 > 0.05.Therefore, we do not reject Null hypothesis and say that the model fits the data well.

- <u>To check Multicollinearity:</u>

GVIF	Df	GVIF^(1/(2*Df))
Y2 2.275176	**4**	**1.108222**
X1 2.571435	**1**	**1.603569**
X2 2.023901	**1**	**1.422639**
X3 2.221589	**1**	**1.490500**
X4 1.803771	**1**	**1.343046**
X5 2.037464	**1**	**1.427398**
X6 1.674401	**1**	**1.293987**
X7 1.848889	**1**	**1.359738**
X8 1.932334	**1**	**1.390084**
X9 1.921527	**1**	**1.386191**

GVIF^(1/(2*df)) close to 1: Indicates little to no multicollinearity.
GVIF^(1/(2*df)) greater than 1: Indicates the presence of multicollinearity, with higher values suggesting more severe multicollinearity.

<u>Interpretation</u>: In the above table all the adjusted GVIF values are close to 1 which indicates relatively little to no collinearity which is acceptable.

- <u>Likelihood ratio test :-</u>

Hypothesis-H0:Reduced model is good as full model
H1: Reduced model is not good as full mode.

Model 1: Y2 ~ X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 + X9
Model 2: Y2 ~ X1 + X2 + X3 + X5 + X6 + X8 + X9

	Df	LogLik	Df	Chisq	Pr(>Chisq)
1	**39**	**-203.26**			
2	**31**	**-209.89**	**-8**	**13.243**	**0.1038**

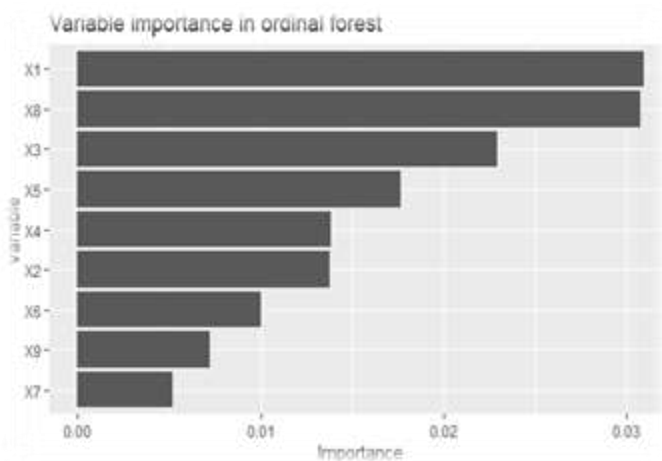
Thus , we do not reject H0 and conclude that Reduced model is good as full model.

- Analysis done Ordinal Forest technique

	Length	Class	Mode
forestfinal	13	ranger	list
bordersbest	6	-none-	numeric
forests	100	-none-	list
perffunctionvalues	100	-none-	numeric
bordersb	600	-none-	numeric
classes	5	-none-	character
nsets	1	-none-	numeric
ntreeperdiv	1	-none-	numeric
ntreefinal	1	-none-	numeric
perffunction	1	-none-	character
classimp	1	-none-	logical
nbest	1	-none-	numeric
classfreq	5	table	numeric
varimp	9	-none-	numeric

Classes of ordinal target variable:
"1", "2", "3", "4", "5"
Importance of individual variable in the ordinal forest model

	variable	Importance
X1	X1	0.030935272
X2	X2	0.013744850
X3	X3	0.022907326
X4	X4	0.013857357
X5	X5	0.017680400
X6	X6	0.010049852
X7	X7	0.005210417
X8	X8	0.030759131
X9	X9	0.007252878



Interpretation-From the table of importance and graph we can see that the variable X1,X8,X3 has the most importance compared to rest of the variables.

- <u>Training accuracy of the above fitted model</u>

  Confusion matrix

pred_labels	1	2	3	4	5
1	3	0	0	0	0
2	1	10	1	0	0
3	0	2	46	1	0
4	0	1	22	58	8
5	0	0	0	4	77
Accuracy=0.8290598					

Accuracy of the training model is **82.90%**

- <u>Testing accuracy of the above fitted model</u>

  Confusion matrix

pred_labels	1	2	3	4	5
1	1	0	0	0	0
2	0	0	0	0	0
3	0	0	12	4	0
4	1	1	7	16	6
5	0	0	0	4	6
Accuracy=0.6034483					

Accuracy of the testing model is **60.34%**

Objective 3

- To study the impact of Artificial Intelligence on Customer Satisfaction

<u>Variable Used:-</u>

- ✓ Outcome Variable (Y) – <u>Customer Satisfaction</u>

<u>Customer Satisfaction</u>:- Customer satisfaction on the impact of AI can be defined as the overall contentment of customers resulting from the utilization of AI technologies to improve their interactions, experiences, and outcomes with a business. This encompasses various dimensions, including service quality, personalization, efficiency, and reliability.

In the objective 3rd the Outcome Variable as well as the Predictor Variables are fitted on Ordinal Scale having following levels –

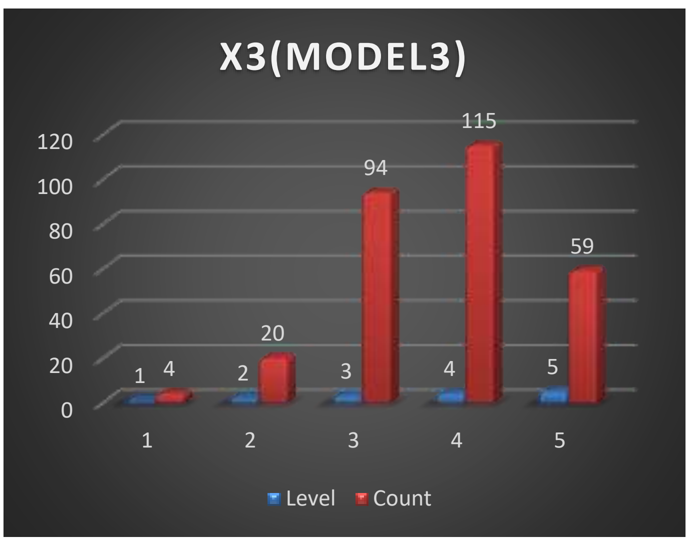
1-Very low
2-Low
3-Moderate
4-High
5-Very high

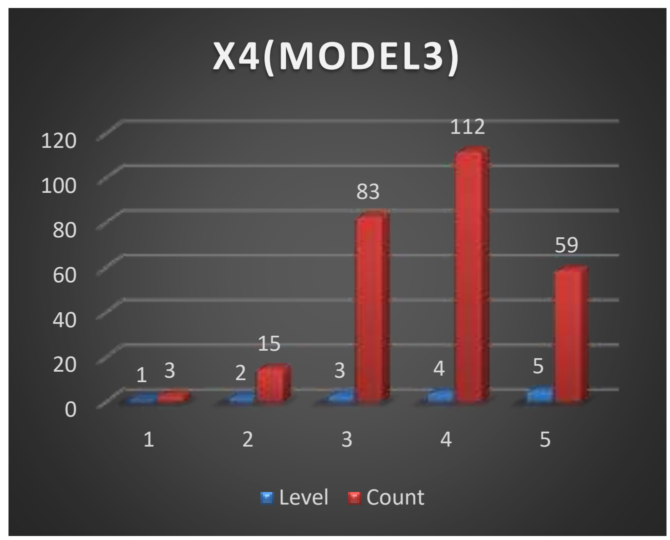
- ✓ <u>Predictor variables</u> –

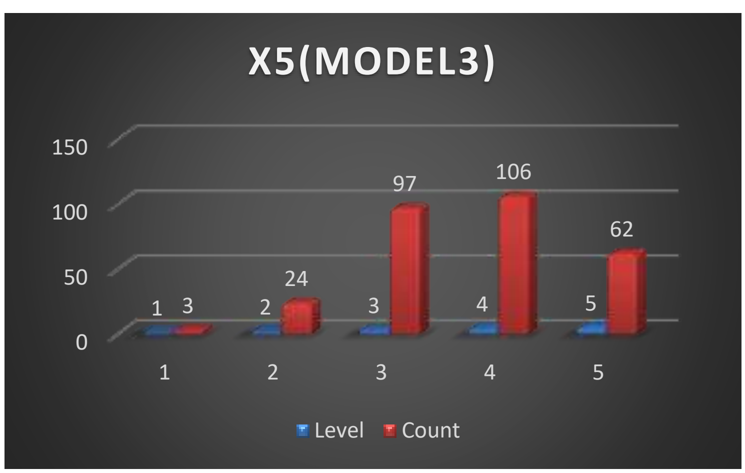
1. **<u>X1- Product quality-</u>** Product Quality refers to the degree to which a product meets certain standards and fulfills customer expectations. It encompasses various attributes such as performance, durability, reliability, aesthetics, and functionality.
2. **<u>X2-Communication with Customer-</u>**Communication with Customer refers to the methods and processes through which a company interacts with its customers to provide information, support, and services. This encompasses various channels such as email, chat, social media, phone calls, and face-to-face interactions.
3. **<u>X3-Accessibility-</u>** Accessibility refers to the design and creation of products, services, and environments that are usable by people with a wide range of abilities and disabilities. It encompasses various aspects, including physical, digital, cognitive, and sensory accessibility.
4. **<u>X4-Timeliness-</u>** Timeliness refers to the extent to which an activity or process is completed within the expected or required timeframe. In the context of AI, timeliness encompasses the ability of AI systems and solutions to accelerate processes, reduce delays, and ensure that tasks and decisions are made promptly and efficiently.

5. **X5-Feedback-** Feedback refers to the information and insights gathered from various sources about the performance, quality, and user experience of a product, service, or system. This information is used to make improvements, correct issues, and enhance overall effectiveness.
6. **X6-Cultural and demographic factors-** Cultural and Demographic Factors refer to the characteristics and behaviors of different groups within a population that influence how they interact with and are affected by artificial intelligence (AI). These factors include age, gender, ethnicity, education, income levels, social values, traditions, and societal norms.
7. **X7-Emotional connection-** Emotional Connection refers to the bond or attachment that individuals feel towards a product, service, brand, or entity. This connection is characterized by feelings of trust, affection, and loyalty.

- <u>Graphical representation of predictors described by levels</u>







Analysis for 3rd Objective

- ❖ (Full model Fitting)
- ❖ Parameter Estimates –

❖ Parameter Estimates								
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Y3 = 1]	-11.827	1.149	106.045	1	<.001	-14.078	-9.576
	[Y3 = 2]	-9.468	.809	136.871	1	<.001	-11.054	-7.882
	[Y3 = 3]	-4.792	.597	64.520	1	<.001	-5.961	-3.623
	[Y3 = 4]	-2.449	.524	21.871	1	<.001	-3.476	-1.423
Location	[X1=1]	-2.430	1.861	1.705	1	.192	-6.077	1.217
	[X1=2]	-2.874	.859	11.193	1	<.001	-4.558	-1.190
	[X1=3]	-1.298	.425	9.301	1	.002	-2.132	-.464
	[X1=4]	-.834	.378	4.862	1	.027	-1.576	-.093
	[X1=5]	0[a]	.	.	0	.	.	.
	[X2=1]	-.092	1.537	.004	1	.952	-3.103	2.920
	[X2=2]	-1.323	.711	3.462	1	.063	-2.717	.071
	[X2=3]	-.332	.446	.555	1	.456	-1.206	.542
	[X2=4]	.239	.376	.405	1	.524	-.498	.976
	[X2=5]	0[a]	.	.	0	.	.	.
	[X3=1]	-1.690	1.229	1.891	1	.169	-4.099	.719
	[X3=2]	-.163	.811	.040	1	.841	-1.752	1.427
	[X3=3]	-.231	.442	.271	1	.602	-1.098	.637
	[X3=4]	.167	.374	.199	1	.655	-.565	.899
	[X3=5]	0[a]	.	.	0	.	.	.
	[X4=1]	3.246	1.463	4.924	1	.026	.379	6.113
	[X4=2]	.158	.835	.036	1	.849	-1.477	1.794
	[X4=3]	-.706	.404	3.061	1	.080	-1.498	.085
	[X4=4]	-.266	.361	.541	1	.462	-.973	.442
	[X4=5]	0[a]	.	.	0	.	.	.
	[X5=1]	-3.399	1.479	5.286	1	.021	-6.297	-.501
	[X5=2]	-2.389	.734	10.582	1	.001	-3.828	-.950
	[X5=3]	-.931	.475	3.836	1	.050	-1.862	.001

	[X5=4]	-.531	.384	1.914	1	.167	-1.283	.221
	[X5=5]	0[a]	.	.	0	.	.	.
	[X6=1]	-2.744	.971	7.991	1	.005	-4.647	-.842
	[X6=2]	-2.799	.691	16.404	1	<.001	-4.153	-1.444
	[X6=3]	-1.684	.466	13.051	1	<.001	-2.598	-.770
	[X6=4]	-1.068	.416	6.601	1	.010	-1.883	-.253
	[X6=5]	0[a]	.	.	0	.	.	.
	[X7=1]	-.992	.670	2.195	1	.138	-2.304	.320
	[X7=2]	-.087	.503	.030	1	.863	-1.072	.899
	[X7=3]	-.734	.424	2.994	1	.084	-1.565	.097
	[X7=4]	-.506	.379	1.783	1	.182	-1.248	.237
	[X7=5]	0[a]	.	.	0	.	.	.

From full model we get 4 variables significant i.e. X1, X4, X5, X6.
X1- Product quality
X4-Timeliness
X5-Feedback
X6-Cultural and demographic factors
(Before fitting the model equation, we go for the reduced model to check whether the reduced model is as good as full model)

- <u>Proportion Odds assumption</u> -

Test of Parallel Lines[a]				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	451.258			
General	406.726[b]	44.532[c]	84	1.000

Hypothesis-H_0: $\beta_{1l} = \beta_1 \quad \forall \quad l = 1, 2, 3, 4$

H_1: $\beta_{1l} \neq \beta_1$ for atleast one l

We Hypothesis that there is common slope parameter for each of the cumulative logit regression equation instead of 4 distinct slopes.

<u>Interpretation</u>-P-value = 1.00> 0.05

Therefore, we do not reject the Null hypothesis and conclude that each outcome variable has a common slope parameter.

- <u>To test Accuracy</u>

Confusion matrix –

Y3 * Predicted Response Category Crosstabulation

Count

		Predicted Response Category					Total
		1	2	3	4	5	
Y3	1	0	0	2	0	0	2
	2	1	1	8	0	0	10
	3	0	0	73	16	8	97
	4	0	0	17	48	22	87
	5	0	0	2	20	74	96
Total		1	1	102	84	104	292

Overall Statistics-Accuracy : 0.6712 , 95% CI : (0.6141, 0.7248), No Information Rate : 0.3322 ,P-Value [Acc > NIR] : < 2.2e-16 , Kappa 0.5172 ,Mcnemar's Test P-Value : NA

Statistics by Class:

	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	0.000000	0.100000	0.7526	0.5517	0.7708
Specificity	0.996552	1.000000	0.8513	0.8244	0.8469
Pos Pred Value	0.000000	1.000000	0.7157	0.5714	0.7115
Neg Pred Value	0.993127	0.969072	0.8737	0.8125	0.8830
Prevalence	0.006849	0.034247	0.3322	0.2979	0.3288
Detection Rate	0.000000	0.003425	0.2500	0.1644	0.2534
Detection Prevalence	0.003425	0.003425	0.3493	0.2877	0.3562
Balanced Accuracy	0.498276	0.550000	0.8019	0.6881	0.8089

The overall accuracy of the full model is <u>67.12%</u> .

❖ (Reduced model fitting)

➢ Parameter estimates

Parameter Estimates							
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval
							Lower Bound
Threshold	[Y3 = 1]	-11.155	1.059	111.009	1	<.001	-13.231
	[Y3 = 2]	-8.911	.737	146.042	1	<.001	-10.356
	[Y3 = 3]	-4.477	.524	72.887	1	<.001	-5.505
	[Y3 = 4]	-2.278	.457	24.856	1	<.001	-3.173
Location	[X1=1]	-1.814	1.685	1.159	1	.282	-5.117
	[X1=2]	-2.666	.734	13.188	1	<.001	-4.104
	[X1=3]	-1.363	.407	11.219	1	<.001	-2.160
	[X1=4]	-.882	.361	5.962	1	.015	-1.589
	[X1=5]	0[a]	.	.	0	.	.
	[X4=1]	2.600	1.289	4.068	1	.044	.074
	[X4=2]	-.429	.732	.343	1	.558	-1.863
	[X4=3]	-.914	.383	5.681	1	.017	-1.665
	[X4=4]	-.252	.345	.536	1	.464	-.927
	[X4=5]	0[a]	.	.	0	.	.
	[X5=1]	-4.100	1.395	8.637	1	.003	-6.834
	[X5=2]	-2.736	.653	17.565	1	<.001	-4.016
	[X5=3]	-1.283	.432	8.803	1	.003	-2.130
	[X5=4]	-.569	.363	2.466	1	.116	-1.280
	[X5=5]	0[a]	.	.	0	.	.
	[X6=1]	-3.416	.881	15.025	1	<.001	-5.144
	[X6=2]	-3.100	.577	28.817	1	<.001	-4.232
	[X6=3]	-1.766	.404	19.076	1	<.001	-2.558
	[X6=4]	-.971	.389	6.222	1	.013	-1.733
	[X6=5]	0[a]	.	.	0	.	.

Fitted Ordinal logistic model-

$$\Pi_1 = \frac{e^{logit\ F1}}{1+e^{logitF1}}$$

Where, **logitF1**=-11.155-1.814*Product Quality(1)-2.666*Product Quality(2)-1.363*Product Quality(3)-0.882*Product Quality(4)+2.600*Timeliness(1)-0.429*Timeliness(2)-0.914*Timeliness(3)-0.252*Timeliness(4)-4.1*Feedback(1)-2.736*Feedback(2)-1.283*Feedback(3)-0.569*Feedback(4)-3.416*Cultural and Demographic Factors(1)-3.100*Cultural and Demographic Factors(2)-1.766*Cultural and Demographic Factors(3)-0.971*Cultural and Demographic Factors(4)

$$\Pi_2 = \frac{e^{logit\ F2}}{1 + e^{logitF2}}$$

Where,**logitF2**=-8.911-1.814*Product Quality(1)-2.666*Product Quality(2)-1.363*Product Quality(3)-0.882*Product Quality(4)+2.600*Timeliness(1)-0.429*Timeliness(2)-0.914*Timeliness(3)-0.252*Timeliness(4)-4.1*Feedback(1)-2.736*Feedback(2)-1.283*Feedback(3)-0.569*Feedback(4)-3.416*Cultural and Demographic Factors(1)-3.100*Cultural and Demographic Factors(2)-1.766*Cultural and Demographic Factors(3)-0.971*Cultural and Demographic Factors(4)

$$\Pi_3 = \frac{e^{logit\ F3}}{1 + e^{logitF3}}$$

Where, **logitF3**=-4.477-1.814*Product Quality(1)-2.666*Product Quality(2)-1.363*Product Quality(3)-0.882*Product Quality(4)+2.600*Timeliness(1)-0.429*Timeliness(2)-0.914*Timeliness(3)-0.252*Timeliness(4)-4.1*Feedback(1)-2.736*Feedback(2)-1.283*Feedback(3)-0.569*Feedback(4)-3.416*Cultural and Demographic Factors(1)-3.100*Cultural and Demographic Factors(2)-1.766*Cultural and Demographic Factors(3)-0.971*Cultural and Demographic Factors(4)

$$\Pi_4 = \frac{e^{logit\ F4}}{1 + e^{logitF4}}$$

Where, **logitF4**=-2.278-1.814*Product Quality(1)-2.666*Product Quality(2)-1.363*Product Quality(3)-0.882*Product Quality(4)+2.600*Timeliness(1)-0.429*Timeliness(2)-0.914*Timeliness(3)-0.252*Timeliness(4)-4.1*Feedback(1)-2.736*Feedback(2)-1.283*Feedback(3)-0.569*Feedback(4)-3.416*Cultural and Demographic Factors(1)-3.100*Cultural and Demographic Factors(2)-1.766*Cultural and Demographic Factors(3)-0.971*Cultural and Demographic Factors(4)

- Proportion Of odds Assumption-

Test of Parallel Lines[a]				
Model	-2 Log Likelihood	Chi-Square	df	**Sig.**
Null Hypothesis	342.970			
General	330.133[b]	12.837[c]	48	**1.000**

We test,

H0: $\beta_{1l} = \beta_1 \; i.e. \; \beta_{11} = \beta_{12} = \beta_{13} = \beta_{14} = \beta_1 \quad \forall \quad l = 1, 2, \dots, k-1$

H1: $\beta_{1l} \neq \beta_1$ for atleast one l

Where β_1 depicts common slope parameter.

We Hypothesis that there is common slope parameter

Interpretation-P-value = 1.00 > 0.05

Therefore, we do not reject Null hypothesis and conclude that there is common slope parameter for each outcome variable.

- Model fitting information and Goodness-of-fit test –

Model Fitting Information				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	565.651			
Final	342.970	222.680	16	<.001

Hypothesis:

H_0: The predictor variables do not significantly improve the fit of the model compared to the null model (without predictors)

H_1: At least one of the predictor variables significantly improves the fit of the model

Interpretation:P-value = 0.001 < 0.05 of Final Model

Therefore, we reject the Null hypothesis and say that the predictor variables significantly improve the fit of the mode.

Goodness-of-Fit			
	Chi-Square	df	Sig.
Pearson	517.408	360	<.001
Deviance	262.181	360	1.000

Hypothesis: H_0: The model fits the data well.

H_1: The model does not fit the data well

Interpretation – P-value – 1.000 >0.05.Therefore, we do not reject Null hypothesis and say that the model fits the data well.

- To Check Multicollinearity –

	GVIF	Df	GVIF^(1/(2*Df))
Y1	2.275176	4	1.108222
X1	2.571435	1	1.603569
X2	2.023901	1	1.422639
X3	2.221589	1	1.490500
X4	1.803771	1	1.343046
X5	2.037464	1	1.427398
X6	1.674401	1	1.293987
X7	1.848889	1	1.359738
X8	1.932334	1	1.390084
X9	1.921527	1	1.386191

GVIF^(1/(2*df)) close to 1: Indicates little to no multicollinearity.

GVIF^(1/(2*df)) greater than 1: Indicates the presence of multicollinearity, with higher values suggesting more severe multicollinearity.

Interpretation: In the above table all the adjusted GVIF values are close to 1 which indicates relatively little to no collinearity which is acceptable.

- <u>To test accuracy</u>

Confusion matrix –

Y3 * Predicted Response Category Crosstabulation

Count

		Predicted Response Category					Total
		1	2	3	4	5	
Y3	1	0	0	2	0	0	2
	2	1	0	9	0	0	10
	3	0	1	68	20	8	97
	4	0	0	17	49	21	87
	5	0	0	1	23	72	96
Total		1	1	97	92	101	292

Overall Statistics - Accuracy : 0.6473 ,95% CI : (0.5895, 0.702) , No Information Rate : 0.3322 , P-Value [Acc > NIR] : < 2.2e-16,Kappa : 0.4827 ,Mcnemar's Test P-Value : NA

Statistics by Class:					
	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	0.000000	0.000000	0.7010	0.5632	0.7500
Specificity	0.996552	0.996454	0.8513	0.7902	0.8520
Pos Pred Value	0.000000	0.000000	0.7010	0.5326	0.7129
Neg Pred Value	0.993127	0.965636	0.8513	0.8100	0.8743
Prevalence	0.006849	0.034247	0.3322	0.2979	0.3288
Detection Rate	0.000000	0.000000	0.2329	0.1678	0.2466
Detection Prevalence	0.003425	0.003425	0.3322	0.3151	0.3459
Balanced Accuracy	0.498276	0.498227	0.7762	0.6767	0.8010

The accuracy of the reduced model is **64.73%**

- <u>Likelihood Ratio Test</u> –

Model 1: Y3 ~ X1 + X2 + X3 + X4 + X5 + X6 + X7
Model 2: Y3 ~ X1 + X4 + X5 + X6

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	44	-212			
2	32	-216.90	-12.	9.1206	0.6926

<u>Hypothesis</u> : H_0:Reduced model is as good as full model

H_1 :Reduced model is not as good as full model

Interpretation: P-value =0.6926>0.05
Hence, we do not reject the null hypothesis and conclude that Reduced model is as good as Full model.

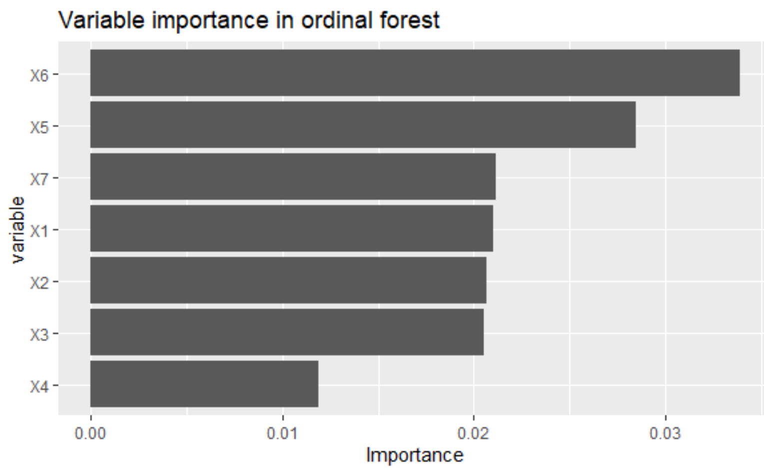
- Analysis done by Ordinal Forest technique

	Length	Class	Mode
forestfinal	13	ranger	list
bordersbest	6	-none-	numeric
forests	100	-none-	list
perffunctionvalues	100	-none-	numeric
bordersb	600	-none-	numeric
classes	5	-none-	character
nsets	1	-none-	numeric
ntreeperdiv	1	-none-	numeric
ntreefinal	1	-none-	numeric
perffunction	1	-none-	character
classimp	1	-none-	logical
nbest	1	-none-	numeric
classfreq	5	table	numeric
varimp	7	-none-	numeric

Classes of ordinal target variable:
"1", "2", "3", "4", "5"

Importance of variables given by ordinal forest technique

	variable	Importance
X1	X1	0.02100203
X2	X2	0.02069795
X3	X3	0.02052550
X4	X4	0.01192263
X5	X5	0.02847894
X6	X6	0.03390493
X7	X7	0.02113547



Interpretation-From the table of importance and graph we can see that the variable X6,X5 has the most importance compared to rest of the variables.

❖ Training accuracy of the above fitted model

Confusion matrix

pred_labels	1	2	3	4	5
1	0	0	0	0	0
2	2	3	2	0	0
3	0	6	78	15	0
4	0	0	0	36	0
5	0	0	8	24	82
Accuracy=**0.7773438**					

Accuracy of the training model is **77.73%**

❖ Testing accuracy of the above fitted model

Confusion matrix

pred_labels	1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	0	0
3	0	1	7	4	1
4	0	0	0	2	0
5	0	0	2	6	13
Accuracy=0.6111111					

Accuracy of the testing model is **61.11%**

❖ Overall Conclusion

From Objective 1-So from the best fitted model ,i.e.reduced model we get 7 variables significant,i.e. X1,X2,X3,X4,X5,X6,X9 which are workforce , capital investment ,Employee productivity, Workplace environment,time management, workload and work life balance, task automation .
From this data analysis, we get know that impact of AI on Efficiency is level 4 ,i.e. HIGH.
Also to increase the accuracy of the model we used machine learning technique ,i.e ordinal forest technique which gives us the accuracy 91.50% on training data and 67.50% on testing data which gives better accuracy ,hence we predict our result from model fitted by ordinal forest technique which gives us the result level 4,i.e HIGH by taking the mode of the predictions.Hence ,the impact of AI on Efficiency of work in industries is HIGH.
From objective 2- From the best fitted model, we get 7 variables significant,i.e. X1, X2, X3, X5, X6, X8, X9 which are Resource management, strategic sourcing, labour cost management, customer relationship management, Supply chain management, Energy efficiency, product design.
From this data analysis, we get know that the impact of AI on Expenditure is level 4 ,i.e. HIGH
From the ordinal forest technique the accuracy is 82.90% on training data and 60.34% on testing data which gives better accuracy, hence we predict our result from the model fitted by ordinal forest technique which gives us the resulting level 4,i.e HIGH by taking the mode of the predictions. Hence,the impact of AI on Expenditure in industries is HIGH.

From objective 3- So from the best-fitted model,i.e.reduced model we get 7 significant variables,i.e. X1,X4,X5,X6 which are product quality, timeliness, feedback, cultural and demographic factors
From this data analysis, we come to know that the impact of AI on Customer Satisfaction is level 5 ,i.e. VERY HIGH
Similarly, objectives 1 & 2 , to increase the accuracy of the model we used machine learning technique,i.e ordinal forest technique which gives us an accuracy of 77.73% on training data and 61.11% on testing data which gives better accuracy,hence we predict our result from the model fitted by ordinal forest technique which gives us the resulting level 5,i.e VERY HIGH by taking the mode of the predictions.
Hence,**the impact of AI on Customer Satisfaction in industries is VERY HIGH.**
Due to not having sufficient data this problem may have been occurred , so , additionally, gathering more data could further improve the model's performance and robustness. For further, use method like cross validation, bagging which can contribute to reducing overfitting by simplifying the model.

Questionnaire :

Name*
Your answer

Age*
Your answer

Gender*
Male
Female
other

Occupation*
Your answer

Type of industry*
Manufacturing
Healthcare
Finance
Retail
Transportation
Energy
Technology
Other:

Does the company currently utilize the Artificial intelligence ?*
Yes
No

How would you rate the AI adoption in your company ?*
Low
Medium
High

For what purpose ,your company is using AI?*
Operations and production
Supply chain management
Customer service and support
Marketing and sales
Research and development
Other:

Which AI applications significantly impact on the efficiency in your organization?*
Predictive Maintainance
Supply chain optimization
Workflow automation
Data analytics
Robotics
Natural Language processing
Computer Vision

How has AI powered data processing and analysis impacted productivity in your organization?*
Faster and accurate decision making

Identifying trends that were previously unnoticed
Improve resource allocation
no significant impact

How has AI driven workflow optimization impacted productivity in your organization?*

Increase efficiency
improved collaboration and communication
Increase output quality
Other:

What are the main challenges company faced while adopting AI ?*

Implementation cost
Lack of expertise and skills
Data privacy and security concerns
Resistance to change
Other:

Impact of AI on following aspects (Efficiency) (swipe to right for other options)*

	Very low	Low	Medium	High	Very high
Efficiency of work					
Workplace environment					
Time management					
Workload and work life balance					
Resource Availability					
Communication and collaboration					
Task automation					
Leadership and management					
Workforce					
Capital investment					
Employee productivity					

Impact of AI on (Cost reduction) (swipe to right for other options)*

	Very low	Low	Medium	High	Very High
Cost savings					
Resource management					
Strategic sourcing					
Labor cost management					
Risk management					
Customer relationship management					
Supply chain management					
Operational efficiency					
Energy efficiency					
Product design					

Impact of AI on (Customer satisfaction) (swipe to right for other options)*

	Very low	Low	Medium	High	Very High
Customer satisfaction					
Product quality					
communication with customer					
Accessibility					
Timeliness					
Feedback					
Cultural and demographic factors					
Emotional connection					

How have AI powered chatbots and virtual assistants impacted customer satisfaction in your organization?*

Enhance accessibilty
Availability of support
Increase customer engagement

❖ R codes:

Objective 1:

```
y=AI_and_industry$`[Efficiency of work]`
x1=AI_and_industry$`[Workforce]`
x2=AI_and_industry$`[Capital investment ]`
x3=AI_and_industry$`[Employee productivity]`
x4=AI_and_industry$`[Workplace environment]`
x5=AI_and_industry$`[Time management]`
x6=AI_and_industry$`[Workload and work life balance]`
x7=AI_and_industry$`[Resource Availability]`
x8=AI_and_industry$`[Communication and collaboration]`
x9=AI_and_industry$`[Task automation]`
x10=AI_and_industry$`[Leadership and management]`
X1=as.factor(na.omit(x1))
X2=as.factor(na.omit(x2))
X3=as.factor(na.omit(x3))
X4=as.factor(na.omit(x4))
X5=as.factor(na.omit(x5))
X6=as.factor(na.omit(x6))
X7=as.factor(na.omit(x7))
X8=as.factor(na.omit(x8))
X9=as.factor(na.omit(x9))
X10=as.factor(na.omit(x10))
Y1=ordered(na.omit(y))
Z1=data.frame(Y1,X1,X2,X3,X4,X5,X6,X7,X8,X9,X10)
library(ordinal)
full_model=clm(Y1~X1+X2+X3+X4+X5+X6+X7+X8+X9+X10,data=Z1)
pred1=predict(full_model,newdata = Z1,type = 'class')
pred2=ordered(as.numeric(unlist(pred1)))
```

```
library(caret)
conf_matrix1=confusionMatrix(pred2,Y1)
conf_matrix1

reduced_model=clm(Y1~X1+X2+X3+X4+X5+X6+X9,data=Z1)
pred3=predict(reduced_model,newdata = Z1,type = 'class')



pred4=ordered(as.numeric(unlist(pred3)))
library(caret)
conf_matrix2=confusionMatrix(pred4,Y1)
conf_matrix2
library(lmtest)
lrt=lrtest(full_model,reduced_model)
lrt

library(dplyr)
data=Z1%>%mutate(X1=as.numeric(X1),X4=as.numeric(X4),X5=as.numeric(X5),X6=as.numeric
(X6),X9=as.numeric(X9))
Y=as.numeric(unlist(Y1))
full_model=lm(Y~.,data = data)
vif_values=vif(full_model)
vif_values
get_mode=function(v){
  uniqv=unique(v)
  uniqv[which.max(tabulate(match(v,uniqv)))]
}
mode=get_mode(pred4)
mode
> y=AI_and_industry$`[Efficiency of work]`
> x1=AI_and_industry$`[Workforce]`
> x2=AI_and_industry$`[Capital investment ]`
> x3=AI_and_industry$`[Employee productivity]`
> x4=AI_and_industry$`[Workplace environment]`
> x5=AI_and_industry$`[Time management]`
> x6=AI_and_industry$`[Workload and work life balance]`
> x7=AI_and_industry$`[Resource Availability]`
> x8=AI_and_industry$`[Communication and collaboration]`
> x9=AI_and_industry$`[Task automation]`
> x10=AI_and_industry$`[Leadership and management]`
> X1=as.factor(na.omit(x1))
> X2=as.factor(na.omit(x2))
> X3=as.factor(na.omit(x3))
> X4=as.factor(na.omit(x4))
> X5=as.factor(na.omit(x5))
> X6=as.factor(na.omit(x6))
> X7=as.factor(na.omit(x7))
> X8=as.factor(na.omit(x8))
> X9=as.factor(na.omit(x9))
> X10=as.factor(na.omit(x10))
> Y1=ordered(na.omit(y))
> Z1=data.frame(Y1,X1,X2,X3,X4,X5,X6,X7,X8,X9,X10)
> library(ggplot2)
> set.seed(1)
> library(caTools)
> sample=sample.split(Z1,SplitRatio = 0.8)
> train=subset(Z1,sample==T)
```

```
> test=subset(Z1,sample==F)
> library(ordinalForest)
> library(ordinalForest)
> model=ordfor(depvar = "Y1",data = train,nsets = 100,ntreeperdiv = 1000)
> summary(model)
> pred=predict(model,newdata=train)
> pred


> pred_labels=pred$ypred
> conf_matrix=table(pred_labels,train$Y1)

> conf_matrix
> Y=as.numeric(unlist(train$Y1))
> pred1=as.numeric(unlist(pred_labels))

> correct=sum(pred_labels==Y)
> accuracy=correct/length(Y)
> accuracy
> set.seed(123)
> importance=model$varimp
> importance_efficiency=data.frame(variable=names(importance),Importance=import
ance)
> importance
> ggplot(importance_efficiency,aes(x=reorder(variable,Importance),y=Importance)
)+geom_bar(stat = "identity")+coord_flip()+xlab("variable")+ylab("Importance")+
ggtitle("Variable importance in ordinal forest")
> pred=predict(model,newdata=test)
> pred
> pred_labels=pred$ypred
> conf_matrix=table(pred_labels,test$Y1)
> conf_matrix
> Y=as.numeric(unlist(test$Y1))
> pred1=as.numeric(unlist(pred_labels))
> correct=sum(pred_labels==Y)
> accuracy=correct/length(Y)
> accuracy
```

Objective2 :

```
y=AI_and_industry$`[Cost savings]`
x1=AI_and_industry$`[Resource management]`
x2=AI_and_industry$`[Strategic sourcing]`
x3=AI_and_industry$`[Labor cost management]`
x4=AI_and_industry$`[Risk  management]`
x5=AI_and_industry$`[Customer relationship management]`
x6=AI_and_industry$`[Supply chain management]`
x7=AI_and_industry$`[Operational efficiency]`
x8=AI_and_industry$`[Energy efficiency]`
x9=AI_and_industry$`[Product design ]`
X1=as.factor(na.omit(x1))
X2=as.factor(na.omit(x2))
X3=as.factor(na.omit(x3))
X4=as.factor(na.omit(x4))
X5=as.factor(na.omit(x5))
X6=as.factor(na.omit(x6))
X7=as.factor(na.omit(x7))
X8=as.factor(na.omit(x8))
X9=as.factor(na.omit(x9))
Y2=ordered(na.omit(y))
Z2=data.frame(Y2,X1,X2,X3,X4,X5,X6,X7,X8,X9)
set.seed(1)
```

```
library(caTools)
sample=sample.split(Z2,SplitRatio = 0.8)
train=subset(Z2,sample==T)
test=subset(Z2,sample==F)
full_model=clm(Y2~X1+X2+X3+X4+X5+X6+X7+X8+X9,data =train)
pred1=predict(full_model,newdata = train,type = 'class')

pred2=ordered(as.numeric(unlist(pred1)))
library(caret)
conf_matrix1=confusionMatrix(pred2,train$Y2)
conf_matrix1
pred3=predict(full_model,newdata = test,type = 'class')

pred4=ordered(as.numeric(unlist(pred3)))
library(caret)
conf_matrix2=confusionMatrix(pred4,test$Y2)
conf_matrix2
reduced_model=clm(Y2~X1+X2+X3+X5+X6+X8+X9,data =train)
pred5=predict(reduced_model,newdata = train,type = 'class')
pred6=ordered(as.numeric(unlist(pred5)))
conf_matrix3=confusionMatrix(pred6,train$Y2)
conf_matrix3
pred7=predict(reduced_model,newdata = test,type = 'class')
pred8=ordered(as.numeric(unlist(pred7)))
conf_matrix4=confusionMatrix(pred8,test$Y2)
conf_matrix4

library(lmtest)
lrt=lrtest(full_model,reduced_model)
lrt

library(car)
vif(full_model)
library(dplyr)
data=Z2%>%mutate(X1=as.numeric(X1),X2=as.numeric(X2),X3=as.numeric(X3),X4=as.numeric
(X4),X5=as.numeric(X5),X6=as.numeric(X6),X7=as.numeric(X7),X8=as.numeric(X8),X9=as.nu
meric(X9))
Y=as.numeric(unlist(Y2))
full_model=lm(Y~.,data = data)
vif_values=vif(full_model)
vif_values
get_mode=function(v){
  uniqv=unique(v)
  uniqv[which.max(tabulate(match(v,uniqv)))]
}
mode1=get_mode(pred4)
mode1
mode2=get_mode(pred8)
mode2
summary(model)
```

```
> pred=predict(model,newdata=train2)
> pred
>pred_labels=pred$ypred
> conf_matrix=table(pred_labels,train2$Y2)
> conf_matrix
> Y=as.numeric(unlist(train2$Y2))
> pred1=as.numeric(unlist(pred_labels))


> correct=sum(pred_labels==Y)
> accuracy=correct/length(Y)
> accuracy

> pred=predict(model,newdata=test2)
> pred

> pred_labels=pred$ypred
> conf_matrix=table(pred_labels,test2$Y2)
> conf_matrix

> Y=as.numeric(unlist(test2$Y2))
> pred1=as.numeric(unlist(pred_labels))
> correct=sum(pred_labels==Y)
> accuracy=correct/length(Y)
> accuracy

> get_mode=function(v){
+   uniqv=unique(v)
+   uniqv[which.max(tabulate(match(v,uniqv)))]
+ }
> pred=predict(model,newdata=train2)
> pred


> pred1=as.numeric(unlist(pred_labels))
> get_mode=function(v){
+   uniqv=unique(v)
+   uniqv[which.max(tabulate(match(v,uniqv)))]
+ }
> mode1=get_mode(pred1)
> mode1

> set.seed(123)
> importance=model$varimp
> importance_cost=data.frame(variable=names(importance),Importance=importance)
> importance
> importance_cost
> library(ggplot2)

> ggplot(importance_cost,aes(x=reorder(variable,Importance),y=Importance))+
geom_bar(stat = "identity")+coord_flip()+xlab("variable")+ylab("Importance")
+ggtitle("Variable importance in ordinal forest")
```

Objective 3:

```
y=AI_and_industry$`[Customer satisfaction]`
x1=AI_and_industry$`[Product quality]`
x2=AI_and_industry$`[ communication with customer]`
x3=AI_and_industry$`[Accessibility]`
x4=AI_and_industry$`[Timeliness]`
x5=AI_and_industry$`[Feedback ]`
x6=AI_and_industry$`[Cultural and demographic factors]`
x7=AI_and_industry$`[Emotional connection]`
X1=as.factor(na.omit(x1))
X2=as.factor(na.omit(x2))
```

```
X3=as.factor(na.omit(x3))
X4=as.factor(na.omit(x4))
X5=as.factor(na.omit(x5))
X6=as.factor(na.omit(x6))
X7=as.factor(na.omit(x7))

Y3=ordered(na.omit(y))
Z3=data.frame(Y3,X1,X2,X3,X4,X5,X6,X7)
library(ordinal)
full_model=clm(Y3~X1+X2+X3+X4+X5+X6+X7,data = Z3)
pred1=predict(full_model,newdata = Z3,type = 'class')
pred2=ordered(as.numeric(unlist(pred1)))
library(caret)
conf_matrix1=confusionMatrix(pred2,Y3)
conf_matrix1

reduced_model=clm(Y3~X1+X4+X5+X6,data = Z3)
pred3=predict(reduced_model1,newdata = Z3,type = 'class')
pred4=ordered(as.numeric(unlist(pred3)))
library(caret)
conf_matrix2=confusionMatrix(pred4,Y3)
conf_matrix2

library(lmtest)
lrt=lrtest(full_model1,reduced_model1)
lrt

library(car)
vif(full_model1)
library(dplyr)
data=Z3%>%mutate(X1=as.numeric(X1),X2=as.numeric(X2),X3=as.numeric(X3),X4=as.numeric
(X4),X5=as.numeric(X5),X6=as.numeric(X6),X7=as.numeric(X7),X8=as.numeric(X8))
Y=as.numeric(unlist(Y3))
full_model=lm(Y~.,data = data)
vif_values=vif(full_model)
vif_values
get_mode=function(v){
  uniqv=unique(v)
  uniqv[which.max(tabulate(match(v,uniqv)))]
}
mode1=get_mode(pred2)
mode1
mode2=get_mode(pred4)
mode2
```

```
> set.seed(1)
> library(caTools)
Warning message:
package ‘caTools’ was built under R version 4.3.2
> sample2=sample.split(Z3,SplitRatio = 0.8)
```

```
> train2=subset(Z3,sample1==T)
> test2=subset(Z3,sample1==F)
> library(ordinalForest)
Warning message:
package ‘ordinalForest’ was built under R version 4.3.3
> model=ordfor(depvar = "Y3",data = train2,nsets = 100,ntreeperdiv = 1000)
> pred=predict(model,newdata=train2)
> pred

"
> pred_labels=pred$ypred
> conf_matrix=table(pred_labels,train2$Y3)
> conf_matrix

> Y=as.numeric(unlist(train2$Y3))
> pred1=as.numeric(unlist(pred_labels))
> correct=sum(pred_labels==Y)
> accuracy=correct/length(Y)
> accuracy
> pred=predict(model,newdata=test2)
> pred

> pred_labels=pred$ypred
> conf_matrix=table(pred_labels,test2$Y3)
> conf_matrix


> Y=as.numeric(unlist(test2$Y3))
> pred1=as.numeric(unlist(pred_labels))
> correct=sum(pred_labels==Y)
> accuracy=correct/length(Y)
> accuracy

> pred=predict(model,newdata=train2)
> pred_labels=pred$ypred
> pred1=as.numeric(unlist(pred_labels))
> get_mode=function(v){
+   uniqv=unique(v)
+   uniqv[which.max(tabulate(match(v,uniqv)))]
+ }
> mode1=get_mode(pred1)
> mode1
> set.seed(123)
> importance=model$varimp
> importance_customer=data.frame(variable=names(importance),Importance=importance)
> importance

> ggplot(importance_customer,aes(x=reorder(variable,Importance),y=Importance))
+geom_bar(stat = "identity")+coord_flip()+xlab("variable")+ylab("Importance")
+ggtitle("Variable importance in ordinal forest")
Error in ggplot(importance_customer, aes(x = reorder(variable, Importance),  :
  could not find function "ggplot"
> library(ggplot2)
> ggplot(importance_customer,aes(x=reorder(variable,Importance),y=Importance))
+geom_bar(stat = "identity")+coord_flip()+xlab("variable")
+ylab("Importance")+ggtitle("Variable importance in ordinal forest")
> summary(model)
```

BIBLIOGRAPHY

References:
1) / https://sqmagazine.co.uk/statistics/artificial-intelligence-statistics
2)https://www.datasciencesociety.net/the-impact-of-artificial-intelligence-on-workplace-efficiency/
3)https://www.valuecoders.com/blog/ai-ml/generative-ai-cost-reduction-strategies/
4)https://www.chitkara.edu.in/blogs/the-impact-of-artificial-intelligence-on-various-industries/