
A Survey on State of Enterprise Machine Learning in Pakistan – 2020

This Survey report is submitted to the Faculty of Computer Science as partial fulfillment of
Master degree of Science in Computer Science

by

Syed Ahmed Hussain Naqvi

Supervised by

Dr. Tariq Mahmood

Professor

Faculty of Computer Science

Institute of Business Administration (IBA), Karachi

Fall Semester 2020

Institute of Business Administration (IBA), Karachi, Pakistan

A Survey on State of Enterprise Machine Learning in Pakistan – 2020

This Survey report is submitted to the Faculty of Computer Science as partial fulfillment of
Master degree of Science in Computer Science

by

Syed Ahmed Hussain Naqvi

Committee:

Dr. Tariq Mahmood

Supervisor

Faculty of Computer Science

Institute of Business Administration (IBA), Karachi

Fall Semester 2020

Institute of Business Administration (IBA), Karachi, Pakistan

Acknowledgement

I would like to thank my supervisor Dr. Tariq Mahmood for guiding me throughout in conducting this research survey. His directions on how to find the right literature and conduct a survey helped me with my literature review and survey formation. He also put forward some great efforts in helping me reach out to the industry participants. IBA Executive Alumni Affairs Department's resources were also utilized to reach out to the participants. In the end, a special thanks to all those authors whose works I have cited in my research survey.

Table of Contents

Abstract	8
1. Introduction.....	9
2. Related Work	10
3. Literature Review	12
A. Machine Learning – Adoption Challenges.....	12
B. Machine Learning – Opportunities and Use cases.....	13
C. Machine Learning – Algorithms and Techniques.....	14
4. Research Methodolgy	17
5. Survey Methodology	19
5. Results.....	21
Use Cases of Machine Learning in Enterprises	21
Human Resource Allocation (%), Budget Allocation (%) and Customer Demand (%) for Machine Learning in Enterprises	24
Adoption Barriers to Machine Learning in Enterprises	26
Algorithms and Learning Techniques for Machine Learning in Enterprises	29
6. Comparative Analysis.....	31
7. Discussion	35
Barriers to Adoption of Machine Learning by Organizational Maturity.....	35
Use Cases of Machine Learning by Organizational Maturity.....	36
Machine Learning Techniques by Organizational Maturity	37
Human Resource Allocation (%), Budget Allocation (%) and Customer Demand (%) for Machine Learning by Organizational Maturity.....	38
Conclusion	40
References	41
Appendix – Research Survey	43

List of Figures

Figure 1: Machine Learning Adoption Barriers, Gartner, 2019 [9].....	12
Figure 2: Use cases of Machine Learning, 2020 State of Enterprise Machine Learning, Algorithmia, 2020 [7].....	13
Figure 3: Classification of Machine Learning Algorithms, A Survey on Machine Learning Approaches and Its Techniques, 2020 [1]	14
Figure 4: Percentage composition of Corporate Sectors in Survey Participants	19
Figure 5: No. of responses over the Survey period.....	20
Figure 6: Survey Results - Use Cases of Machine Learning in Enterprises in Pakistan, 2020	21
Figure 7: Survey Results – Sector-wise Use Cases of Machine Learning in Enterprises in Pakistan, 2020.....	22
Figure 8: Survey Results - Maturity Levels of Machine Learning in Enterprises in Pakistan, 2020	23
Figure 9: Survey Results - Human Resource Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020	24
Figure 10: Survey Results - Budget Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020	24
Figure 11: Survey Results - Customer Demand (%) for Machine Learning in Enterprises in Pakistan, 2020	25
Figure 12: Survey Results - Adoption Barriers to Machine Learning in Enterprises in Pakistan, 2020.....	26
Figure 13: Survey Results – Sector-wise Adoption Barriers to Machine Learning in Enterprises in Pakistan, 2020	27
Figure 14: Survey Results - Machine Learning Techniques used in Enterprises in Pakistan, 2020	29
Figure 15: Survey Results - Machine Learning Algorithms used (Grouping based on similarity) in Enterprises in Pakistan, 2020.....	30

Figure 16: Comparison of Maturity Level of Machine Learning in Enterprise Survey Results in Pakistan - 2020 with Algorithmia Survey Results - 2020 [7].....	31
Figure 17: Comparison of Adoption Barriers to Machine Learning in Enterprise Survey Results in Pakistan - 2020 with Gartner Survey Results - 2019 [9].....	32
Figure 18: Comparison of Machine Learning Use cases in Enterprise Survey Results in Pakistan - 2020 with Algorithmia Survey Results - 2020 [7]	33
Figure 19: Survey Results – Organizational Maturity-wise Adoption Barriers to Machine Learning in Enterprises in Pakistan, 2020.....	35
Figure 20: Survey Results – Organizational Maturity-wise Use Cases of Machine Learning in Enterprises in Pakistan, 2020.....	36
Figure 21: Survey Results – Organizational Maturity-wise Machine Learning Techniques used in Enterprises in Pakistan, 2020.....	37
Figure 22: Survey Results – Organizational Maturity-wise Human Resource Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020	38
Figure 23: Survey Results – Organizational Maturity-wise Budget Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020.....	39
Figure 24: Survey Results – Organizational Maturity-wise Customer Demand (%) for Machine Learning in Enterprises in Pakistan, 2020.....	39

List of Tables

Table 1: Classification of Machine Learning Algorithms by their similarity, A Tour of Machine Learning Algorithms by Jason Brownlee, 2019 [2].....	16
Table 2: Research survey sources and their contributions (%).....	18
Table 3: Number of articles at each step of SLR methodology	18

Abstract

This paper focuses on evaluating the state of Machine Learning in Enterprises in Pakistan in 2020. It highlights the Machine Learning use cases in Pakistan and the challenges in its adoption in business solutions. A survey was conducted in enterprises in Pakistan ranging from companies with well-developed machine learning lifecycles to those with little or no experience in machine learning. The results of this survey have been shared in this paper along with comparisons with earlier works in this field.

Keywords—Machine Learning, Regression Algorithms, Instance-based Algorithms, Regularization Algorithms, Decision Tree Algorithms, Bayesian Algorithms, Clustering Algorithms, Association Rule Learning Algorithms, Boosting Algorithms, Artificial Neural Network Algorithms, Deep Learning Algorithms

1. Introduction

With the tremendous amounts of data at hand and availability of high compute powers, Machine Learning has become extremely popular because of its capabilities. Unlike traditional approaches, it analyses the data and constructs models on them. These models can then be used to make decisions or predictions. Also they can give insights which can be helpful in understanding the data.

Machine Learning has a wide range of applications such as natural language processing (NLP), pattern recognition, fraud detection, robotics, data mining, recommender systems, sales prediction etc. In all these applications, it usually deals with following major problems: clustering, regression and classification.

The learning techniques involved in Machine Learning can be broadly classified into unsupervised learning, semi-supervised learning, supervised learning and reinforcement learning [1]. Besides this, Machine Learning algorithms can also be grouped by similarity in their design and functionality. These groupings include Regression Algorithms, Instance-based Algorithms, Regularization Algorithms, Decision Tree Algorithms, Bayesian Algorithms, Clustering Algorithms, Association Rule Learning Algorithms, Boosting Algorithms, Artificial Neural Network Algorithms, and Deep Learning Algorithms etc. [2]

This survey aims to evaluate the state of Machine Learning in Enterprises in Pakistan in 2020. This is because, at present, Pakistan has a population of around 217 million people with around 114 million youngsters which are under the age of 25. According to the Octo-verse GitHub Community Report [3], Pakistan ranks 6th among the fastest-growing countries of open source projects. Also, more and more companies in Pakistan are getting involved in this domain.

2. Related Work

Over the years, many surveys have been conducted around the world to find out the applications of Machine Learning, its adoption barriers and its impact on different sectors of life such as medical services, financial markets [4], arts and sports etc. More recently some surveys have also been conducted in Pakistan along similar lines.

In 2019, a python notebook [5] was published on Kaggle using data from 2019 Kaggle Machine Learning & Data Science Survey [6]. The aim of this work was to provide an in-depth view of the State of Machine Learning in Pakistan. The key findings of this work are as follows:

1. Majority of participants were male and students
2. Most participants were Undergraduates
3. Google Cloud Platform is chosen more often than Amazon Web Services and Microsoft Azure
4. Most of the Machine Learning work is in small companies
5. The most preferred programming language for Machine Learning is Python
6. The most favorite media sources for Machine Learning are Kaggle and Youtube
7. Most of the participants working in Machine Learning had less than 1 year and 1-2 year experience
8. Most of the Machine Learning work is in exploration phase and have not been used in production [5]

In 2020, Algorithmia published a white paper [7] on its website building on the State of Enterprise Machine Learning report that they published in 2018. In this survey, thousands of companies from around the globe with different maturity levels in Machine Learning were contacted. About 750 companies across all industries participated in the survey. The participants were divided into 2 groups; those which were sent the survey by a third-party and those which were sent the survey by Algorithmia. The key findings of this survey are as follows:

1. Companies often have less than 10 employees with the role of data scientist but this number is growing

2. Most common use cases for Machine Learning are customer-centric but more varied business use cases are also coming up
3. Deployment process of Machine Learning is still immature but it is maturing with time
4. Scale, version control, model reproducibility, and aligning stakeholders are the main challenges in development of Machine Learning
5. It takes between 31 and 90 days to deploy a model for most companies
6. Machine Learning budget allocation is growing most often by 25 percent, and the banking, manufacturing, and IT industries have seen the largest budget growth this year
7. Machine Learning success is being determined by organizations by both business unit and statistical metrics with a significant divide by job level [7]

3. Literature Review

A. Machine Learning – Adoption Challenges

Adoption of Machine Learning has faced many challenges in different parts of the world. For most businesses, the biggest challenge for adoption of Machine Learning is the skills of staff. Developers in Machine Learning domain are often young and have less than 2 years job experience. Besides, for a successful Machine Learning project, experts from different fields are required while companies often rely the knowledge of data engineers [8] [9]. Another big challenge is data scope and quality. Organizations working in Machine Learning need to have data management strategies, data sanity checks and data governance mechanisms [8] [9].

Apart from this, most businesses fail to understand the benefits and use cases of Machine Learning. Organizations need to identify some key areas which are financially feasible and then invest in those areas only [8]. Also, Organizations must develop a clear understanding of what Machine Learning can and cannot do [9].

Moreover, like other technologies, Machine Learning also raises security and privacy concerns. At present, usually large businesses are collecting customer data to increase their profits rather than sharing it for the good of the people [10]. As a result, enough data is not available for using Machine Learning to its full capacity. Fig. 1 below shows some of the adoption barriers of Machine Learning found in a Gartner Survey conducted in 2019 [9].

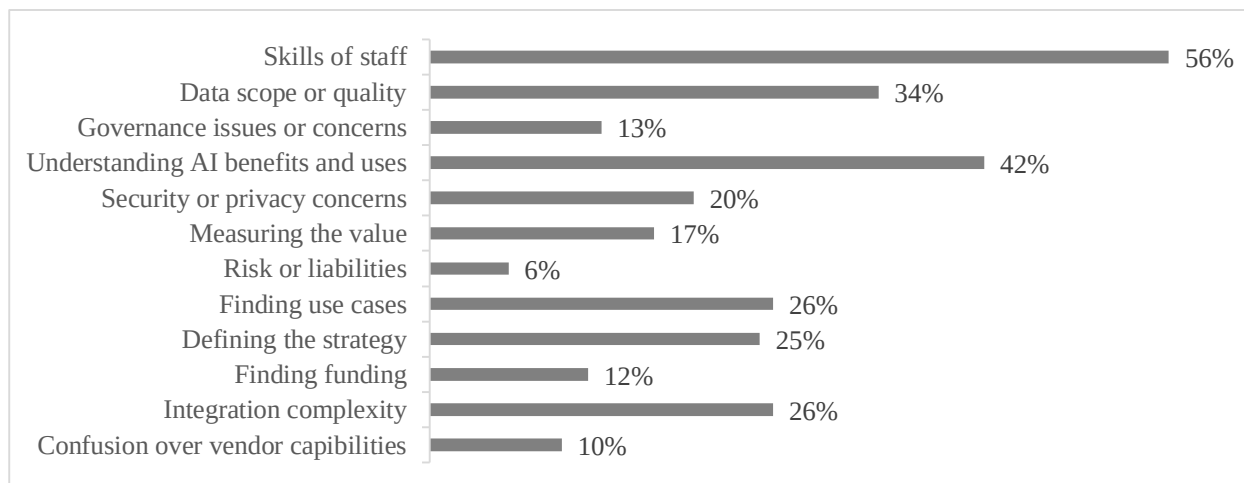


Figure 1: Machine Learning Adoption Barriers, Gartner, 2019 [9]

B. Machine Learning – Opportunities and Use cases

There has been a rise in adoption of Machine Learning over the years. The Stamford firm's 2019 CIO Survey of 89 countries and more than 3,000 executives in [11] shows that there has been a growth of 270 percent since 2015. It is estimated that, by 2022, this market will be worth 6.14 billion dollars. A big factor in this growth is the way its concepts has been aligned between academia, scientific researches and corporates [8]. According to Gartner Research [12], organizations are more interested in using Machine Learning for improving customer experience by interacting with them and increasing their satisfaction, loyalty and engagement etc. and automating tasks and processes. Moreover, personalized marketing and product recommendations can be done by using the behavioral data of customers [13].

Predictions and inferential assertions like fraud detection is another research field in Machine Learning [13] [14]. It is often used in banks for the safety of consumers and saves money spent on disputes and chargebacks.

In [13], Algorithmia also highlights some other uses cases of Machine Learning like Dynamic pricing, Email filtering, Voice assistants etc. Fig. 2 shows some of the use cases of Machine Learning found in a survey by Algorithmia conducted in 2020 [7].

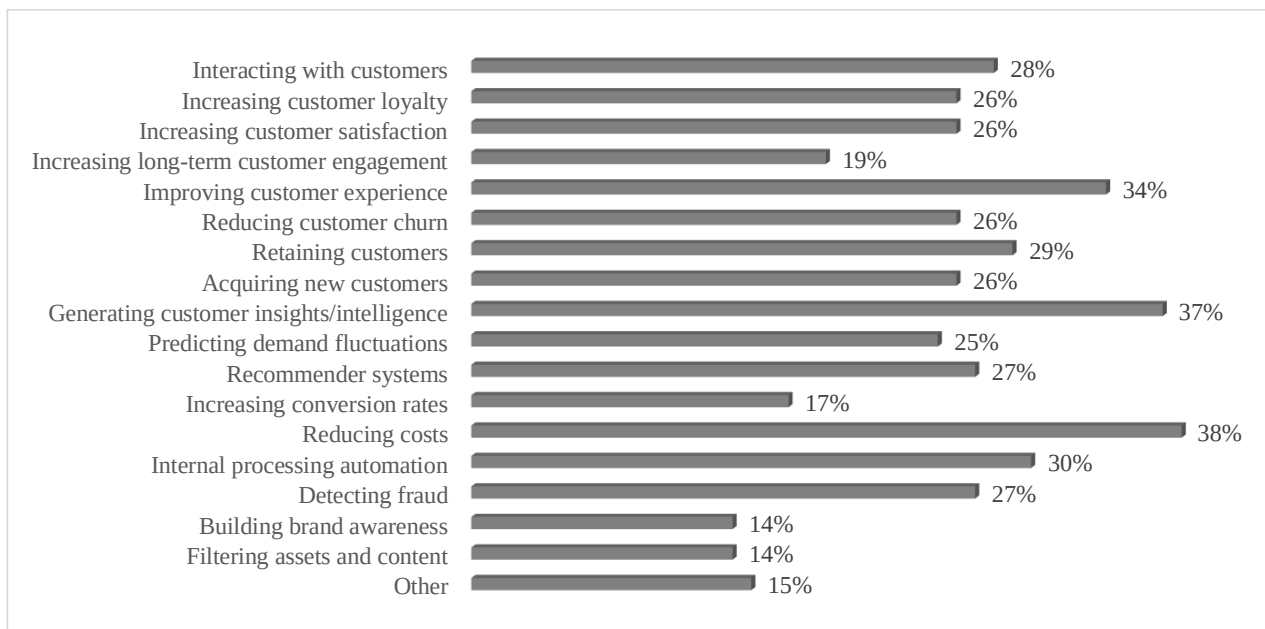


Figure 2: Use cases of Machine Learning, 2020 State of Enterprise Machine Learning, Algorithmia, 2020 [7]

C. Machine Learning – Algorithms and Techniques

One way of grouping Machine Learning Algorithms is to group them by their learning style [2]. Fig. 3 shows how this classification can be done [1].

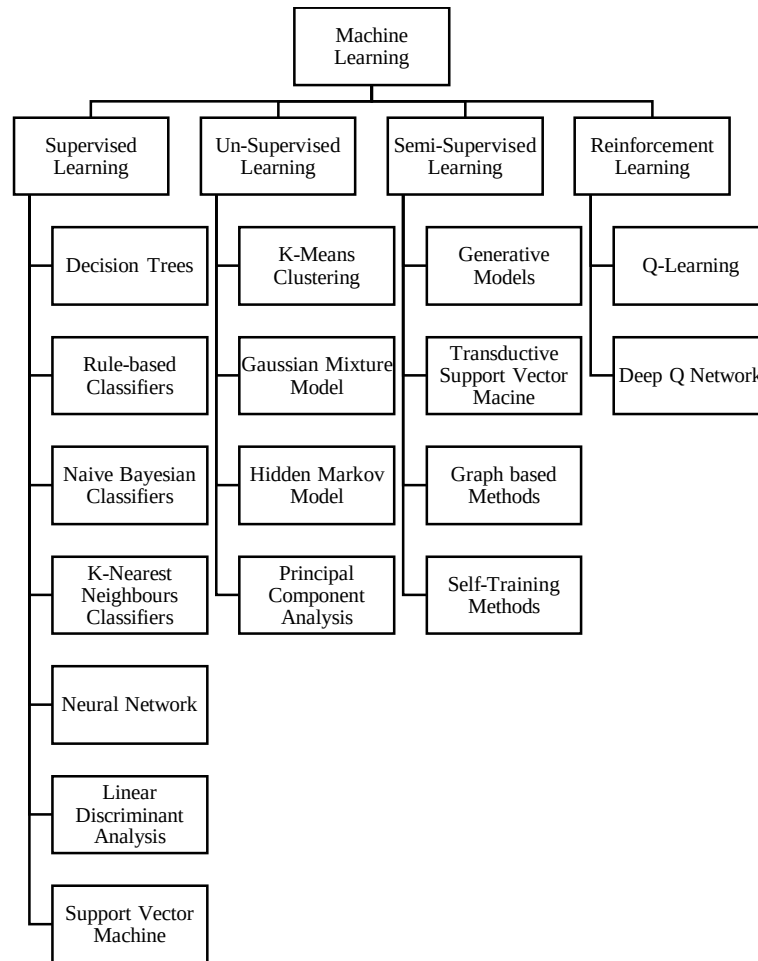


Figure 3: Classification of Machine Learning Algorithms, A Survey on Machine Learning Approaches and Its Techniques, 2020 [1]

There are four different learning styles in Machine Learning Algorithms, namely:

1. Supervised Learning:

In this learning style, the input data is labelled and training is done on the data to produce a model. In this training, the model is required to make predictions. If the predictions are wrong the model is corrected. The process continues until a desired level of accuracy is

achieved by the model on the training data. Example problems are classification and regression. [2]

2. Unsupervised Learning:

In this learning style, the input data is not labeled and structures are deduced in the data through a mathematical process to produce a model and to extract general rules. This can help reduce redundancy and organize data by similarity. Example problems are clustering, dimensionality reduction and association rule learning. [2]

3. Semi-Supervised Learning:

In this learning style, the input data is a mixture of partially labeled and partially unlabelled. Training is done to make predictions but the model must also deduce the structures in the data in order to make these predictions. Example problems are classification and regression. [2]

4. Reinforcement Learning:

In this learning style, the software agent gathers from the interaction with the environment to take actions that would maximize the reward [1]. There is positive reward for each correct action and a penalty for each incorrect decision. It is suitable in cases when you have little data about the problem.

Another useful way of grouping Machine Learning Algorithms is to group them by their similarity in form or function [2]. Although many algorithms can fit into multiple groups, some of these groupings are shown in Table 1.

Groups	Algorithms
Regression Algorithms	Ordinary Least Squares Regression (OLSR), Linear Regression, Logistic Regression, Stepwise Regression, Multivariate Adaptive Regression Splines (MARS), Locally Estimated Scatterplot Smoothing (LOESS)
Instance-based Algorithms	k-Nearest Neighbor (kNN), Learning Vector Quantization (LVQ), Self-Organizing Map (SOM), Locally Weighted Learning (LWL), Support Vector Machines (SVM)
Regularization Algorithms	Ridge Regression, Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net, Least-Angle Regression (LARS)
Decision Tree Algorithms	Classification and Regression Tree (CART), Iterative Dichotomiser 3 (ID3), C4.5 and C5.0 (different versions of a powerful approach), Chi-squared Automatic Interaction Detection (CHAID), Decision Stump, M5, Conditional Decision Trees
Bayesian Algorithms	Naive Bayes, Gaussian Naive Bayes, Multinomial Naive Bayes, Averaged One-Dependence Estimators (AODE), Bayesian Belief Network (BBN), Bayesian Network (BN)
Clustering Algorithms	k-Means, k-Medians, Expectation Maximization (EM), Hierarchical Clustering
Association Rule Learning Algorithms	Apriori Algorithm, Eclat Algorithm
Artificial Neural Network Algorithms	Perceptron, Multilayer Perceptrons (MLP), Back-Propagation, Stochastic Gradient Descent, Hopfield Network, Radial Basis Function Network (RBFN)
Deep Learning Algorithms	Convolutional Neural Network (CNN), Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), Stacked Auto-Encoders, Deep Boltzmann Machine (DBM), Deep Belief Networks (DBN)
Dimensionality Reduction Algorithms	Principal Component Analysis (PCA), Principal Component Regression (PCR), Partial Least Squares Regression (PLSR), Sammon Mapping, Multidimensional Scaling (MDS), Projection Pursuit, Linear Discriminant Analysis (LDA), Mixture Discriminant Analysis (MDA), Quadratic Discriminant Analysis (QDA), Flexible Discriminant Analysis (FDA)
Ensemble Algorithms	Boosting, Bootstrapped Aggregation (Bagging), AdaBoost, Weighted Average (Blending), Stacked Generalization (Stacking), Gradient Boosting Machines (GBM), Gradient Boosted Regression Trees (GBRT), Random Forest

Table 1: Classification of Machine Learning Algorithms by their similarity, A Tour of Machine Learning Algorithms by Jason Brownlee, 2019 [2]

4. Research Methodology

The research conducted for this survey was carried out using a systematic literature review (SLR) methodology. The methodology was used for collecting data related to research questions and summarizing them accordingly. Following steps were taken in carrying out this methodology:

1. **Domain Research:** The domain selected for this research survey is Machine Learning. In this step, the authors went through all the recent publications that have been made in this domain. This literature review helped in formulating the questions for the research survey.
2. **Problem Statement Research:** The problem statement selected for this research survey is finding out the state of machine learning in enterprises in Pakistan in 2020. In this step, the authors went through all the surveys that have been conducted in this domain especially those that have been conducted in Pakistan. Some of these works have already been discussed in detail in section 2 of this paper.
3. **Research using Search Queries:** In this step, initially the authors created search queries according to the research title. Some of these search queries are as follows:
 - 1) Machine Learning in Enterprises
 - 2) Machine Learning in Pakistan
 - 3) State of Machine Learning
 - 4) Machine Learning 2020
 - 5) Machine Learning in Industry

In the next phase, the authors created search queries according to the research questions. Some of these search queries are as follows:

- 1) Use Cases of Machine Learning
- 2) Maturity Level of Machine Learning
- 3) Human Resource Allocation for Machine Learning
- 4) Budget Allocation for Machine Learning
- 5) Customer Demand for Machine Learning
- 6) Adoption Barriers to Machine Learning
- 7) Challenges of Machine Learning

- 8) Machine Learning Algorithms
- 9) Machine Learning Algorithm Types
- 10) Machine Learning Techniques used in Enterprises

These search queries were used for research in several digital sources. These sources include IEEE, ACM, Wiley, Elsevier, Springer, Google Scholar and Blogs. The percentage contribution of all these sources in the articles found at this stage shown in Table 2 below.

S No.	Sources	Contribution (%)
1	IEEE	40
2	ACM	25
3	Wiley	2
4	Elsevier	3
5	Springer	3
6	Google Scholar	20
7	Blogs	7

Table 2: Research survey sources and their contributions (%)

The results of these search queries were filtered to be from years 2015 to present.

4. **Filtration using Title, Abstract and Text Content:** In this step, the articles collected were filtered on the basis of Research paper Title and then Abstract and then based on the text contents of the paper. The number of articles which were selected after these filtrations were applied are shown in Table 3 below.

S No.	Step	Number of articles
1	Search queries	100
2	Title Filtration	85
3	Abstract Filtration	30
4	Text Content Filtration	20

Table 3: Number of articles at each step of SLR methodology

In this way, all the relevant articles were found and cited in this research paper.

5. Survey Methodology

In recent times, there has been a significant progress in adoption of Machine Learning in Enterprises in Pakistan. This survey was conducted in December 2020 and January 2021. In this survey, more than 300 Enterprises in Pakistan were contacted.

These enterprises belonged to different corporate sectors of Pakistan like Education, Financial / Professional Services, Manufacturing, Textiles, Oil and Gas production, Insurance, Utilities etc. The percentage representation of different sectors in survey participants are shown in Fig. 4 below.

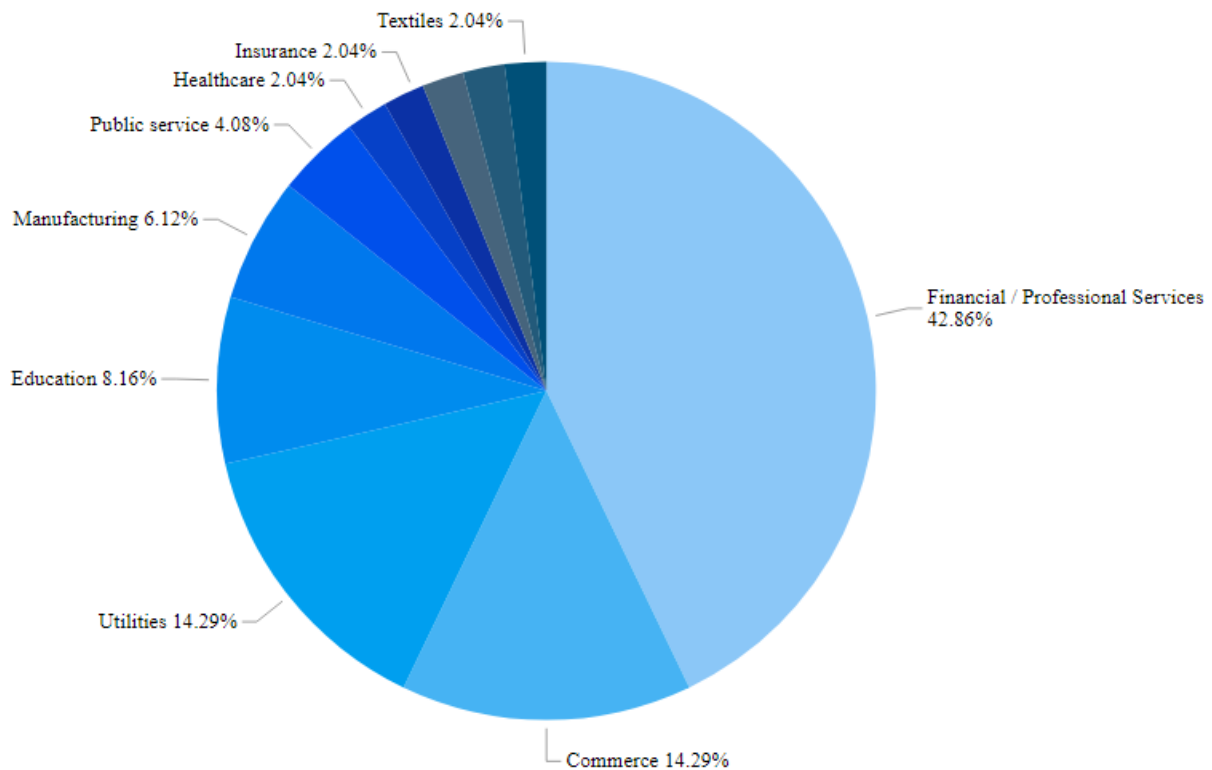


Figure 4: Percentage composition of Corporate Sectors in Survey Participants

Responses were received from the participants as shown in Fig.5 below.

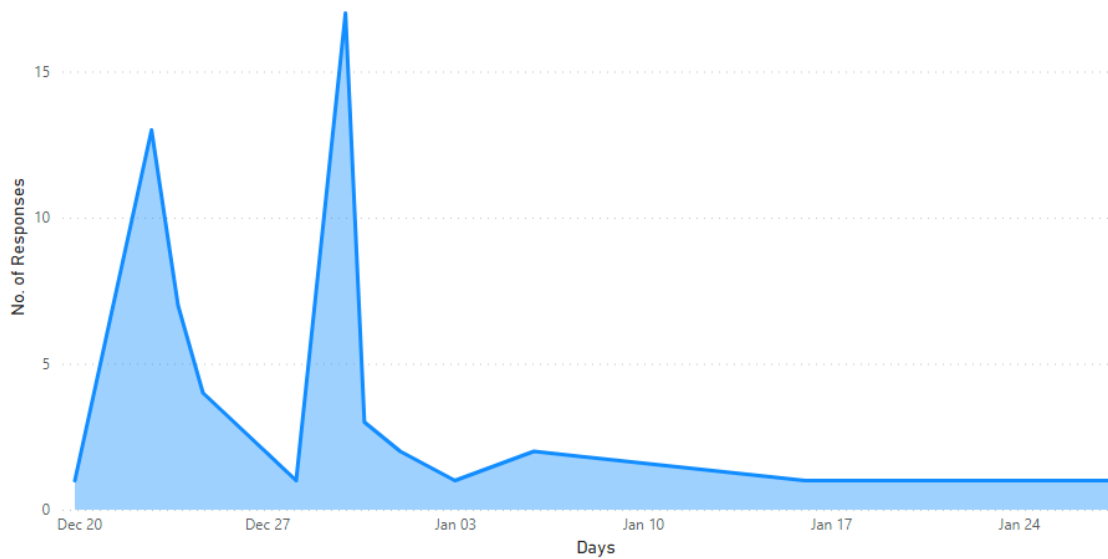


Figure 5: No. of responses over the Survey period

There were initial spikes in responses in early days of survey period as more and more companies were contacted. Later on, companies continued to respond as and when they were sent reminders. In this survey, we have focused on following questions related to state of Machine Learning in Enterprises in Pakistan:

1. Use Cases of Machine Learning in Enterprises
2. Maturity Level of Machine Learning in Enterprises
3. Human Resource Allocation (%) for Machine Learning in Enterprises
4. Budget Allocation (%) for Machine Learning in Enterprises
5. Customer Demand (%) for Machine Learning in Enterprises
6. Adoption Barriers to Machine Learning in Enterprises
7. Machine Learning Algorithms used (Grouping based on similarity) in Enterprises
8. Machine Learning Techniques used in Enterprises

If the participants gave vague answers like “I don’t know” or “Not related to us”, these answers were removed from the results. Also, responses from participants which did not disclose their names were also removed from results.

Results

In this section, we will discuss the findings of this survey along with their graphical representations.

Use Cases of Machine Learning in Enterprises

In this survey, we provided a list of use cases of Machine Learning to the companies to choose from. Participants were allowed to choose all the use cases that they have working on. This list was inspired from [7] so that we can compare the survey results with its findings. Our survey results are shown in Fig. 6 below.

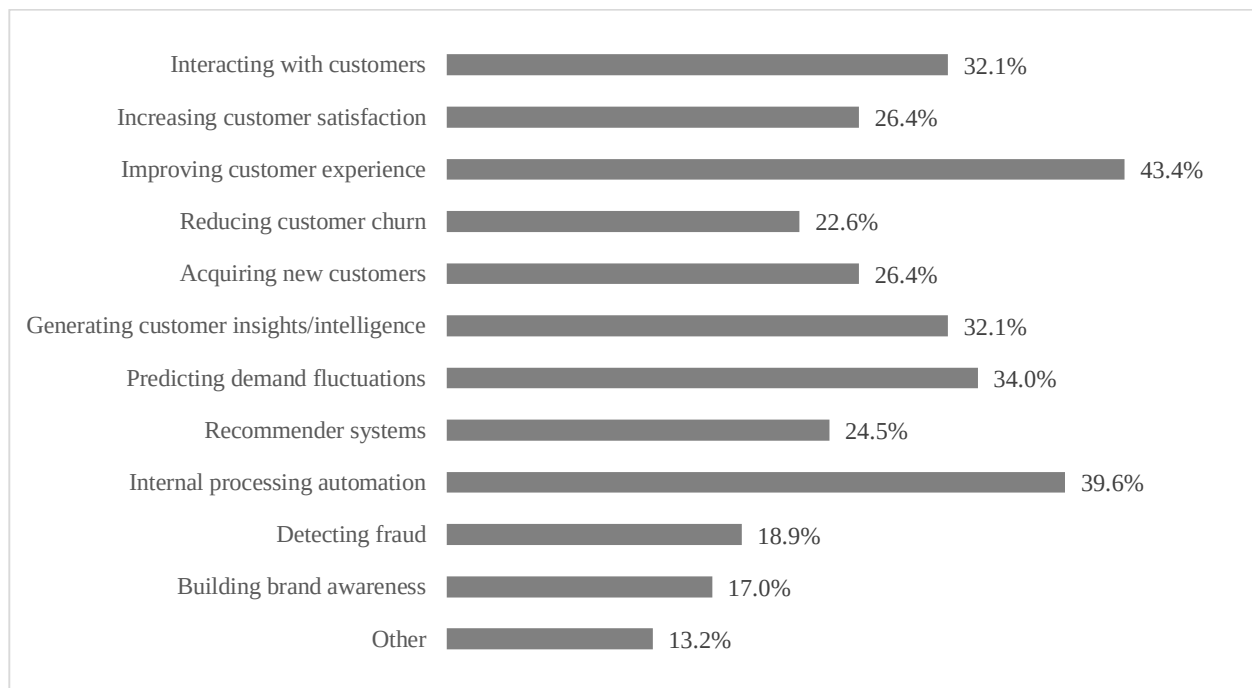


Figure 6: Survey Results - Use Cases of Machine Learning in Enterprises in Pakistan, 2020

Overall, according to the survey results, the top 5 Machine Learning Use Cases in Pakistan are:

1. Improving Customer Experience (43.4%)
2. Internal Process Automations (39.6%)
3. Predicting Demand Fluctuations (34.0%)
4. Generating Customer Insights / Intelligence (32.1%)
5. Interacting with Customers (32.1%)

If we breakdown the use cases with respect to the industry sector of participants, the results of the survey are as shown in Fig. 7. Following are the key findings of these results:

1. Majority of use cases of Machine Learning have been explored in Financial / Professional Services sector (Banks, Finance Companies, IT companies, etc.), Commerce and Educational institutions. Utilities and Textile industry have also found some use cases mostly related to customer experience and demand fluctuations.
2. The most common use case across all the sectors is predicting demand fluctuations which indicates that most of the sectors in Pakistan are focusing on using Machine Learning to catch up with the customer demands to either prevent losses or gain competitive advantage.
3. Financial / Professional Services sector (Banks, Finance Companies, IT companies, etc.), Commerce and Public Services are focusing on fraud detection and internal process automation which shows that these might be serious challenges for them.
4. In the Healthcare sector, generating customer insights and recommender systems are most popular use cases. This shows that a big challenge for this sector is lack of awareness about the patients and making well-informed decisions.
5. Improving customer experience is a major focus of most of the sectors that provide products and services to the customers like Textiles, Utilities, Manufacturing, Commerce and Financial Services etc.

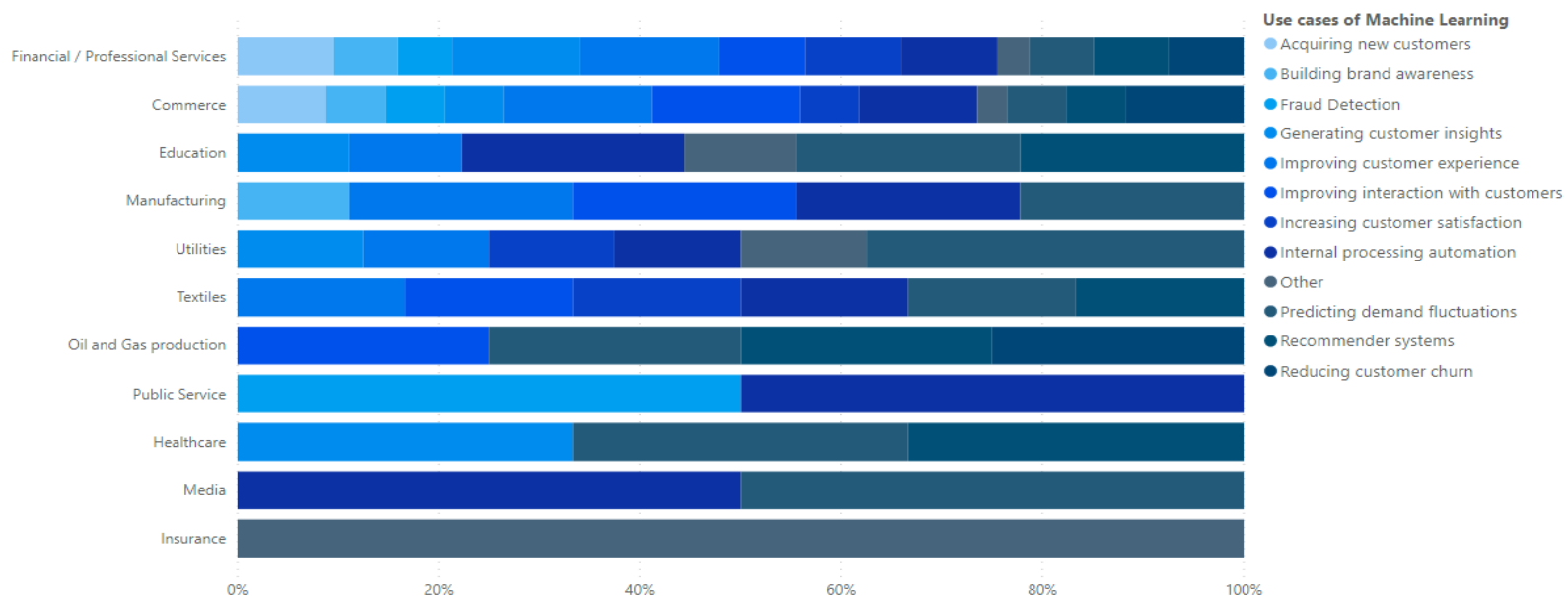


Figure 7: Survey Results – Sector-wise Use Cases of Machine Learning in Enterprises in Pakistan, 2020

Maturity Level of Machine Learning in Enterprises

In this survey, we asked the participant enterprises to gauge how mature is their organization in the field of Machine Learning. Our survey results are shown in Fig. 8 below.

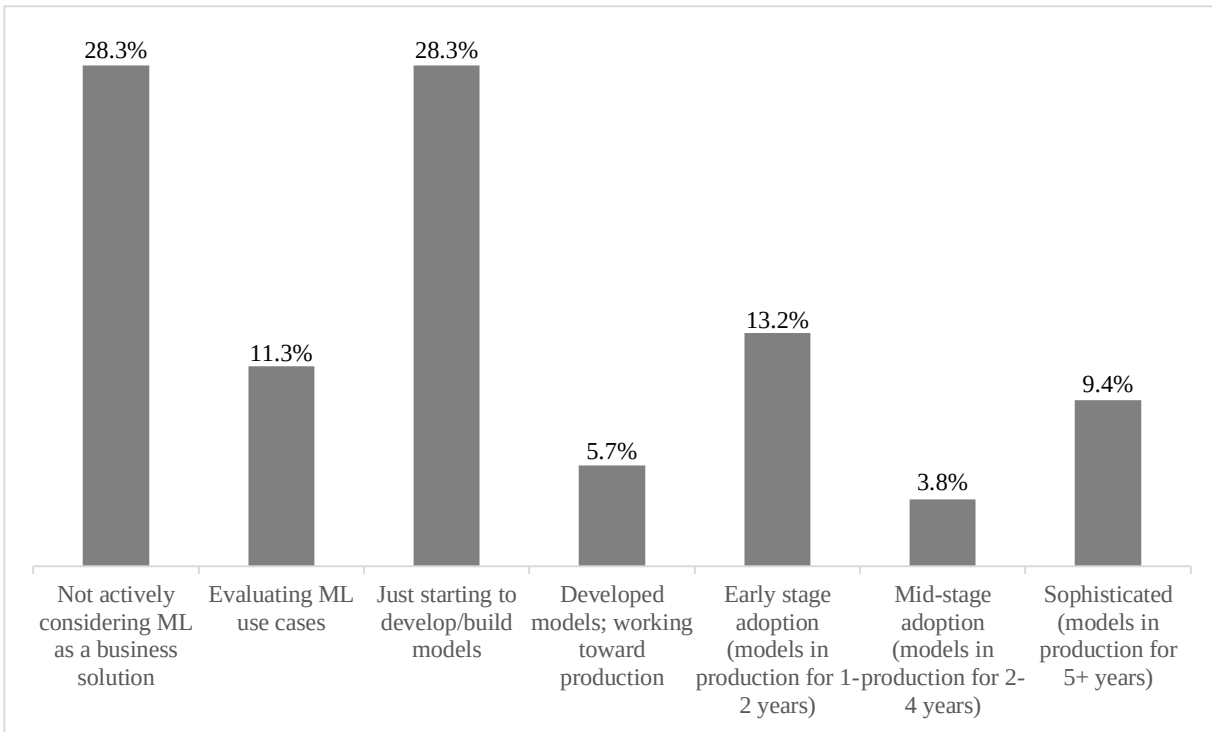


Figure 8: Survey Results - Maturity Levels of Machine Learning in Enterprises in Pakistan, 2020

The key findings of these survey results are as follows:

1. About 28.3% of the enterprises surveyed are not actively considering Machine Learning as a business solution.
2. About 73.6% of the enterprises surveyed have not deployed any Machine Learning model in production.
3. Majority of the enterprises who have deployed any Machine Learning model in production are in early stage adoption (models in production for 1-2 years).
4. About 9.4% of the enterprises surveyed have models in production for more than 5 years.

This is a good sign for Pakistan which is gradually maturing in this field.

Human Resource Allocation (%), Budget Allocation (%) and Customer Demand (%) for Machine Learning in Enterprises

In this survey, we have also looked into the Human Resource Allocation (%), Budget Allocation (%) and Customer Demand (%) for Machine Learning in enterprises in Pakistan. Our survey results are shown in Figs. 9, 10 and 11 below.

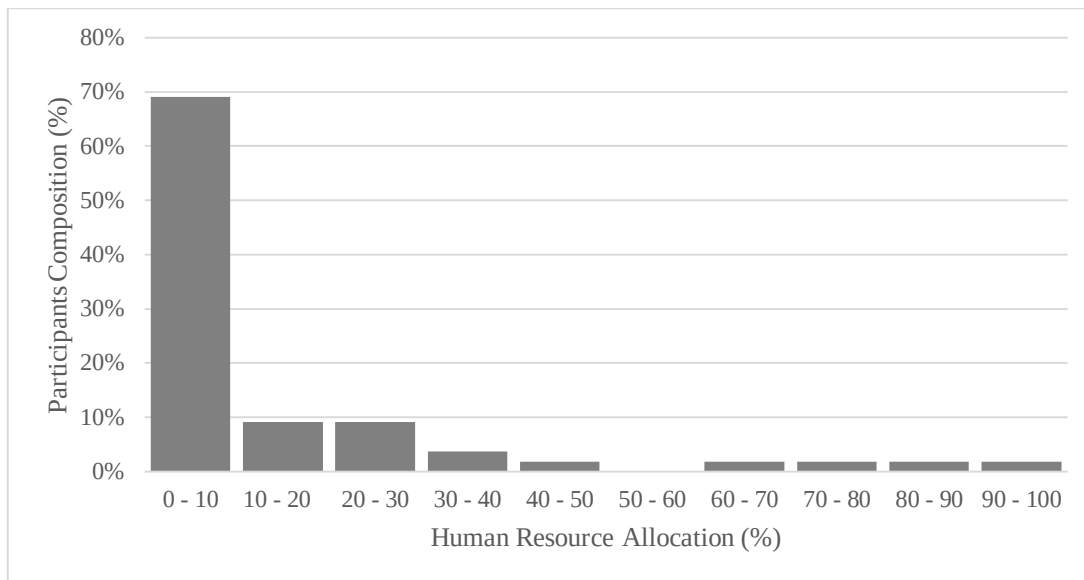


Figure 9: Survey Results - Human Resource Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020

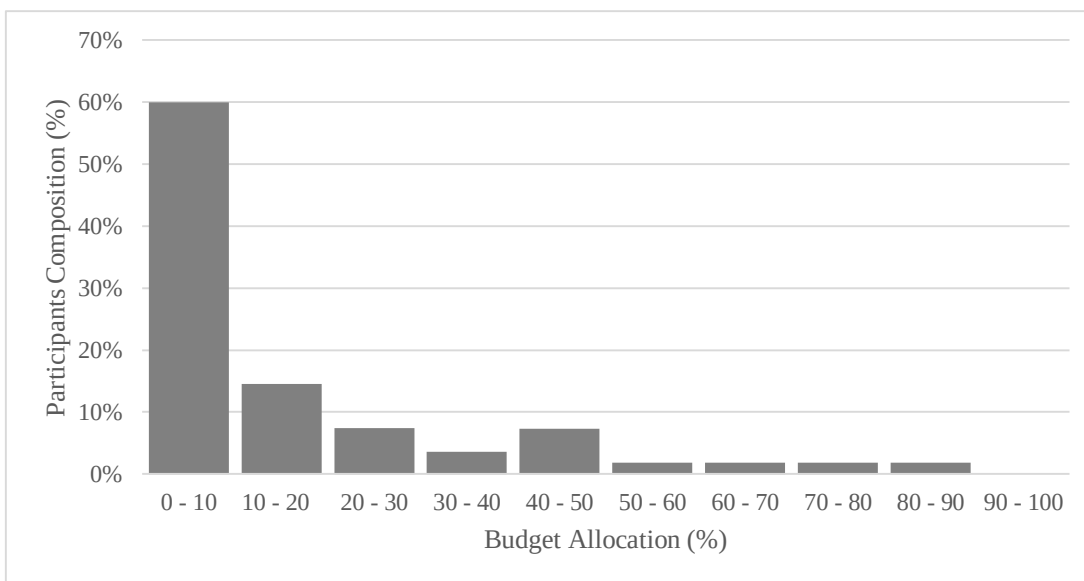


Figure 10: Survey Results - Budget Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020

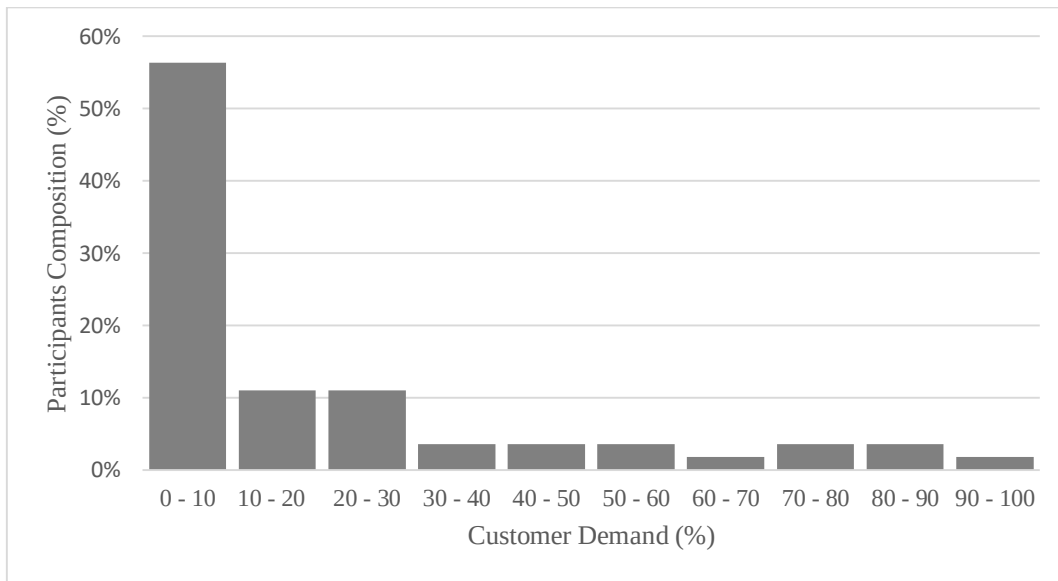


Figure 11: Survey Results - Customer Demand (%) for Machine Learning in Enterprises in Pakistan, 2020

The key findings of these survey results are as follows:

1. About 69% of the enterprises surveyed have about 0-10% Human Resource Allocation for Machine Learning
2. About 60% of the enterprises surveyed have about 0-10% Budget Allocation for Machine Learning
3. About 57% of the enterprises surveyed have about 0-10% of customers demanding for Machine Learning

Adoption Barriers to Machine Learning in Enterprises

In this survey, we provided a list of adoption barriers to Machine Learning to the companies to choose from. Participants were allowed to choose all the barriers that they have faced in the adoption of Machine Learning. This list was inspired from [9] so that we can compare the survey results with its findings. Our survey results are shown in Fig. 12 below.

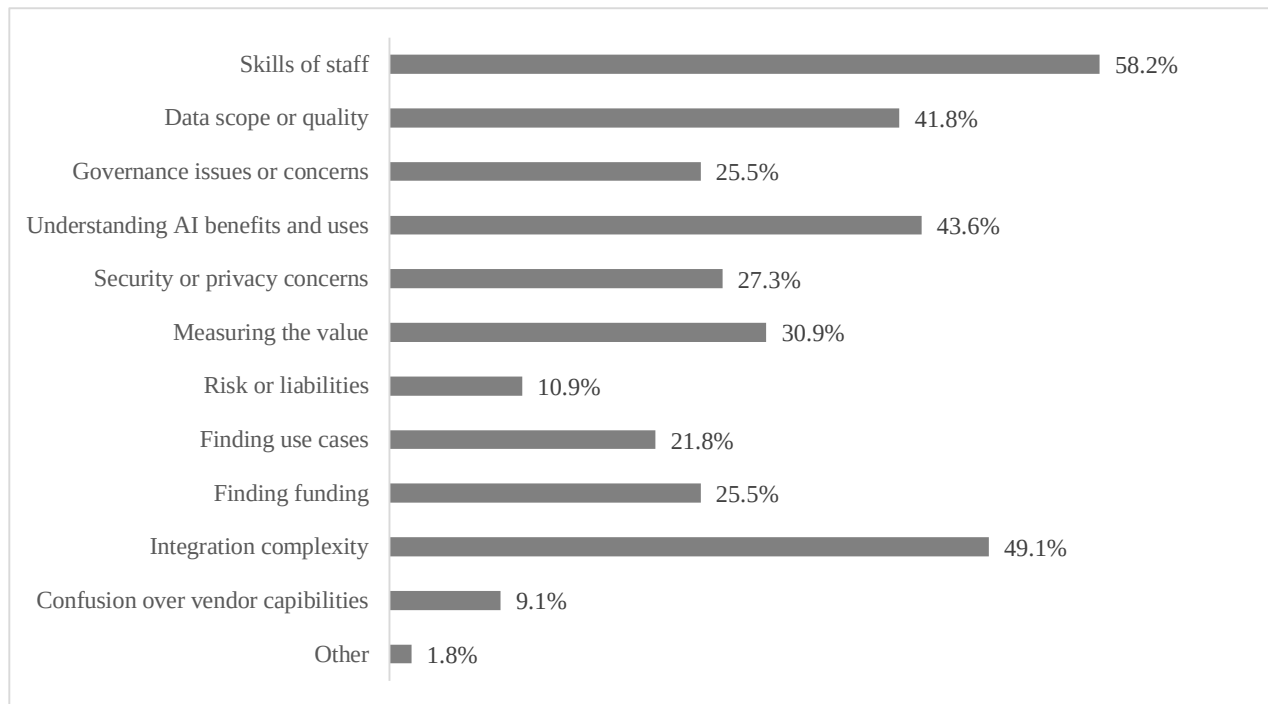


Figure 12: Survey Results - Adoption Barriers to Machine Learning in Enterprises in Pakistan, 2020

Overall, according to the survey results, the top 5 Barriers to adoption of Machine Learning in Pakistan are:

1. Skills of staff (58.2%)
2. Integration complexity (49.1%)
3. Understanding AI benefits and uses (43.6%)
4. Data scope or quality (41.8%)
5. Measuring the value (30.9%)

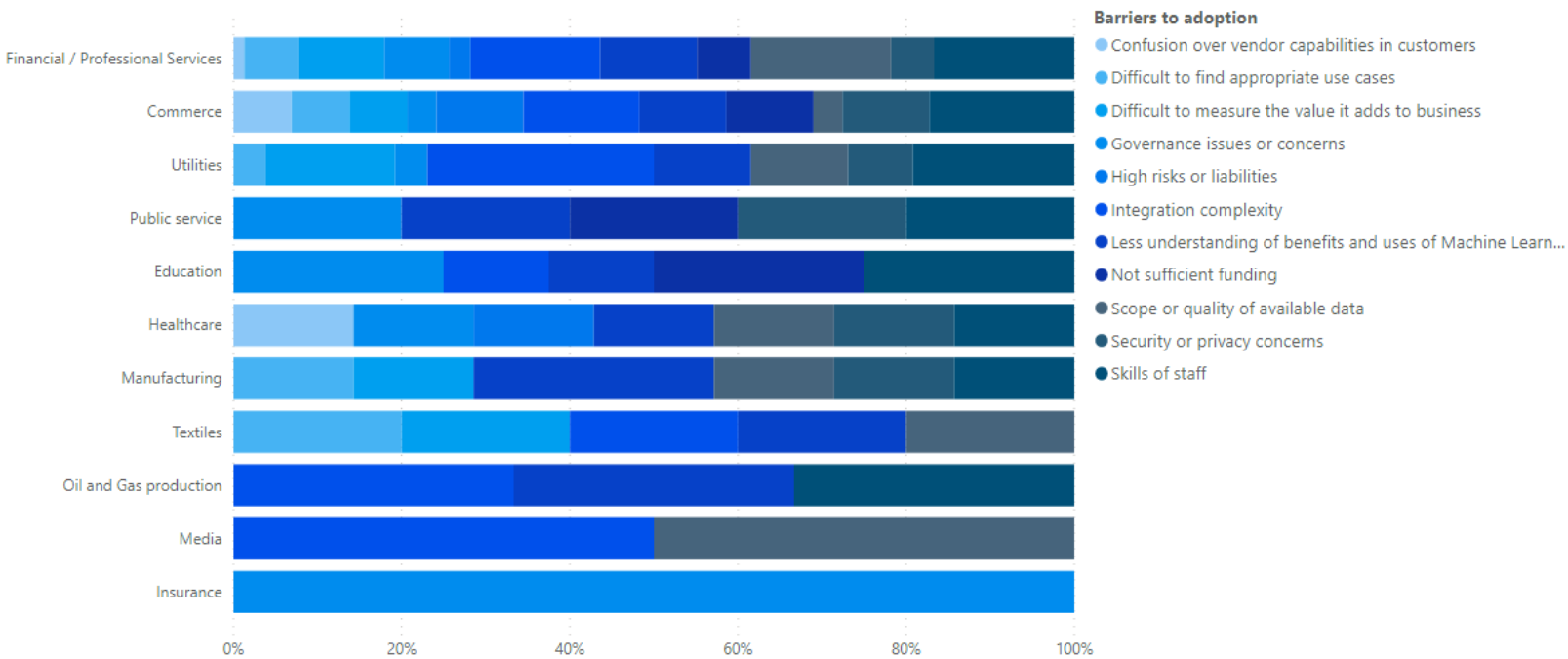


Figure 13: Survey Results – Sector-wise Adoption Barriers to Machine Learning in Enterprises in Pakistan, 2020

If we break down the barriers to adoption of Machine Learning with respect to the industry sector of participants, the results of the survey are as shown in Fig. 13 above. Following are the key findings of these results:

1. Majority of adoption barriers of Machine Learning have been identified by Financial / Professional Services (Banks, Finance Companies, IT companies, etc.), Commerce and Utilities sectors. Healthcare and Education institutions have also identified some barriers to its adoption.
2. The most common adoption barriers across all the sectors are skills of staff and less understanding of benefits and uses of Machine Learning. This indicates that much effort is needed, both in academia and industry of Pakistan, on producing skilled Machine Learning engineers and experts.
3. Industry sectors that provide products and services to the customers like Textiles, Utilities, Manufacturing, Commerce and Financial Services etc. have responded that it is difficult to measure the value Machine Learning adds to the business. Thus, these sectors need Machine Learning experts and analysts to guide them in this direction.

4. In the Healthcare sector, there are also adoption barriers like scope or quality of data, governance issues and confusion over vendor capabilities. This shows that this sector requires efforts in establishing data governance mechanisms in collaboration with more reliable vendors.
5. Security or privacy concerns have been identified as an adoption barrier by 6 major industrial sectors but it has not been identified as an adoption barrier by the educational institutions. This shows that educational institutions need to focus more on these concerns.

Algorithms and Learning Techniques for Machine Learning in Enterprises

Since Machine Learning involves using a learning technique, we asked the survey participants to choose which of the techniques shown in Fig. 14 below they have utilized in their organizations.

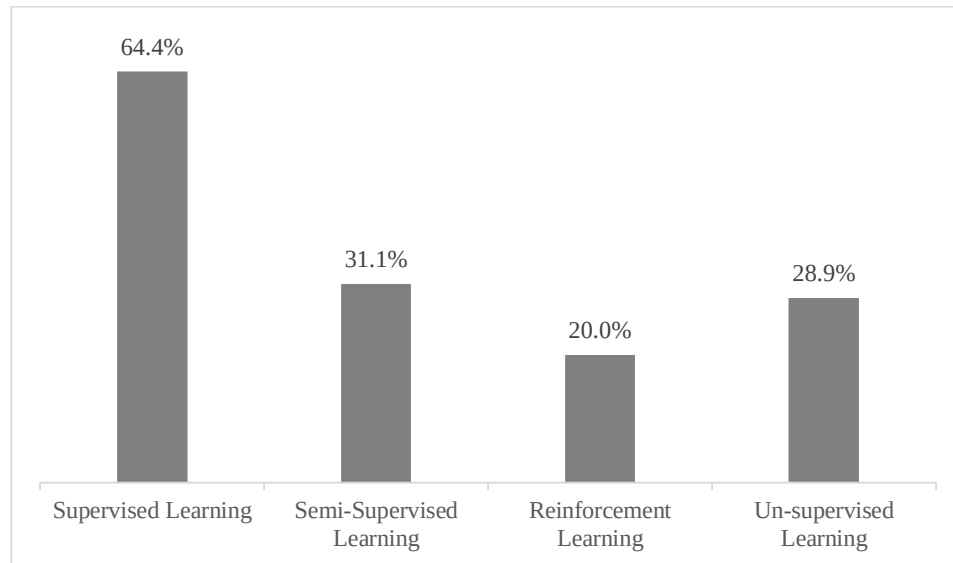


Figure 14: Survey Results - Machine Learning Techniques used in Enterprises in Pakistan, 2020

The key findings of the survey results are as follows:

1. The most popular learning technique is Supervised Learning (64.4%)
2. Semi-supervised Learning (31.1%) and Un-supervised Learning (28.9%) are also gaining popularity amongst the enterprises
3. Reinforcement Learning (20.0%) is also happening which is a good sign as it is a big challenge to prepare the environment for this kind of learning.

Besides this, Machine Learning algorithms can also be grouped on their similarity. Fig. 15 shows the results of which algorithm groups are utilized in enterprises in Pakistan.

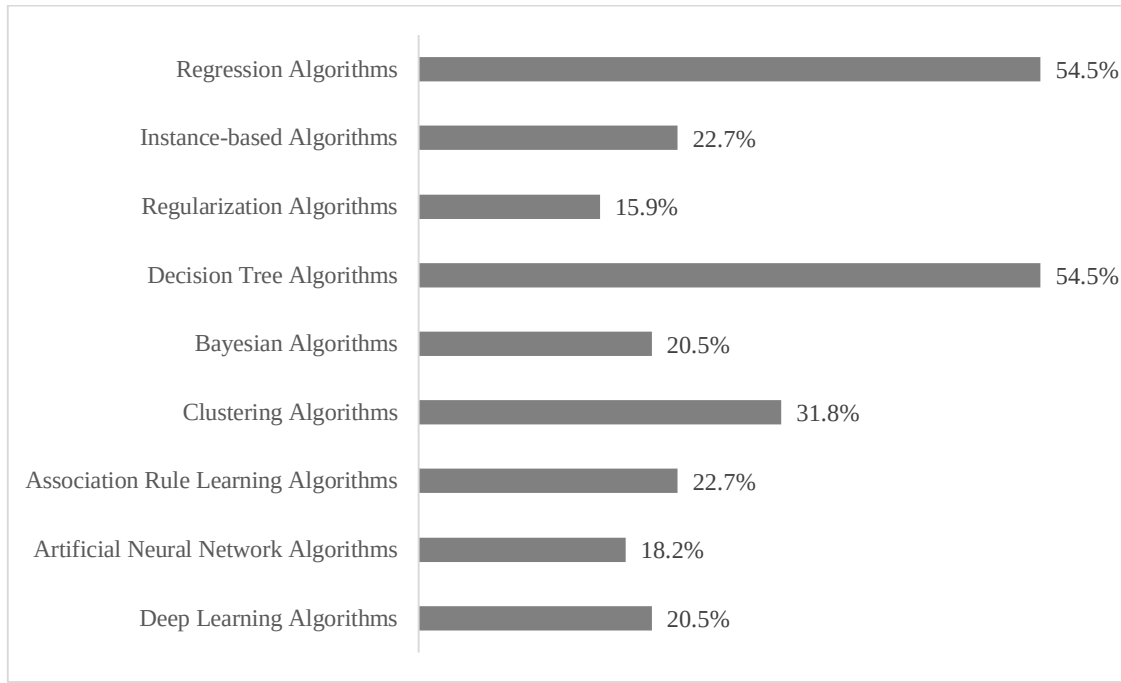


Figure 15: Survey Results - Machine Learning Algorithms used (Grouping based on similarity) in Enterprises in Pakistan, 2020

Overall, according to the survey results, the top 5 Machine Learning Algorithm groups used in Pakistan are:

1. Regression Algorithms (54.5%)
2. Decision Tree Algorithms (54.5%)
3. Clustering Algorithms (31.8%)
4. Instance-based Algorithms (22.7%)
5. Association Rule Learning Algorithms (32.1%)

6. Comparative Analysis

In this section, we will compare our survey results with those in [7] and [9]. This analysis can help us put the survey results of Pakistan into a perspective and compare it with results around the globe.

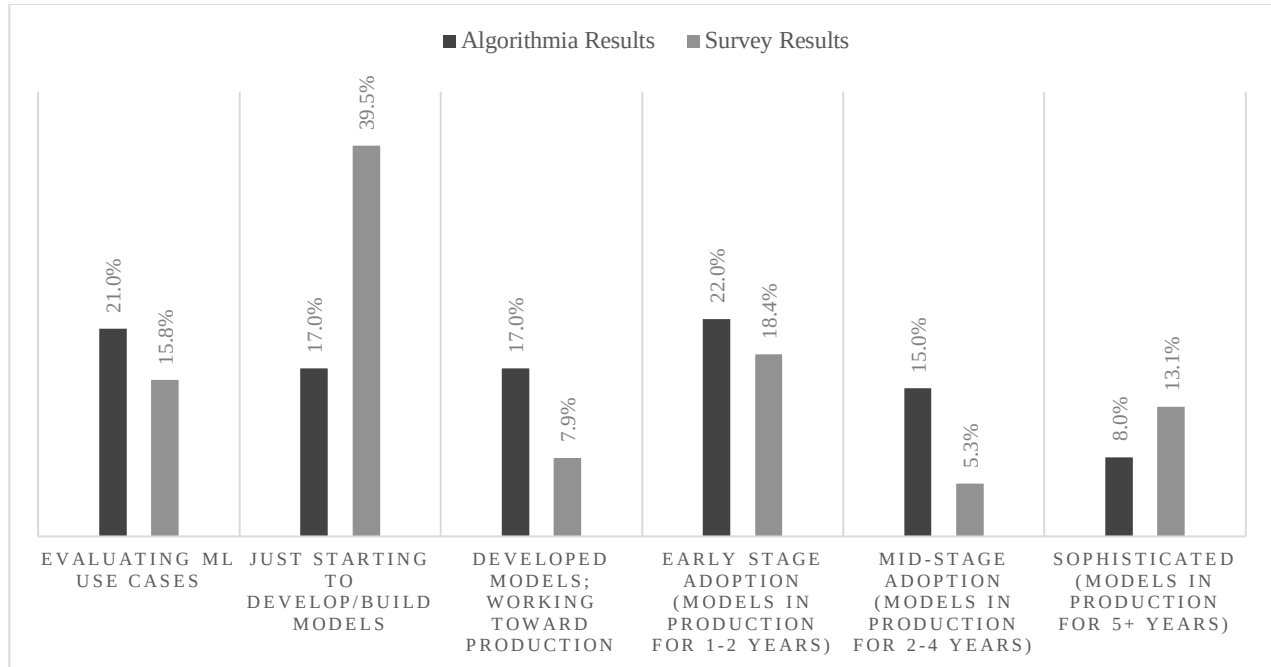


Figure 16: Comparison of Maturity Level of Machine Learning in Enterprise Survey Results in Pakistan - 2020 with Algorithmia Survey Results - 2020 [7]

Fig. 16 shows the results for maturity levels of organizations that participated in our survey and the survey conducted by Algorithmia [7]. In this figure, we have removed the results of participants who are not actively considering Machine Learning as a business solution. From those who are considering it as a business solution, about 63.2% of the enterprises surveyed in Pakistan have not deployed any Machine Learning model in production. This is a much bigger percentage compared to 55% in Algorithmia survey results in [7]. The percentage of the enterprises which are in early stages of development and working towards production is also much less in Pakistan's survey results i.e. 7.9% compared to 17% in Algorithmia survey results in [7]. Moreover, only 36.8% of the enterprises surveyed in Pakistan have any model deployed in production compared to 45% in Algorithmia survey results in [7].

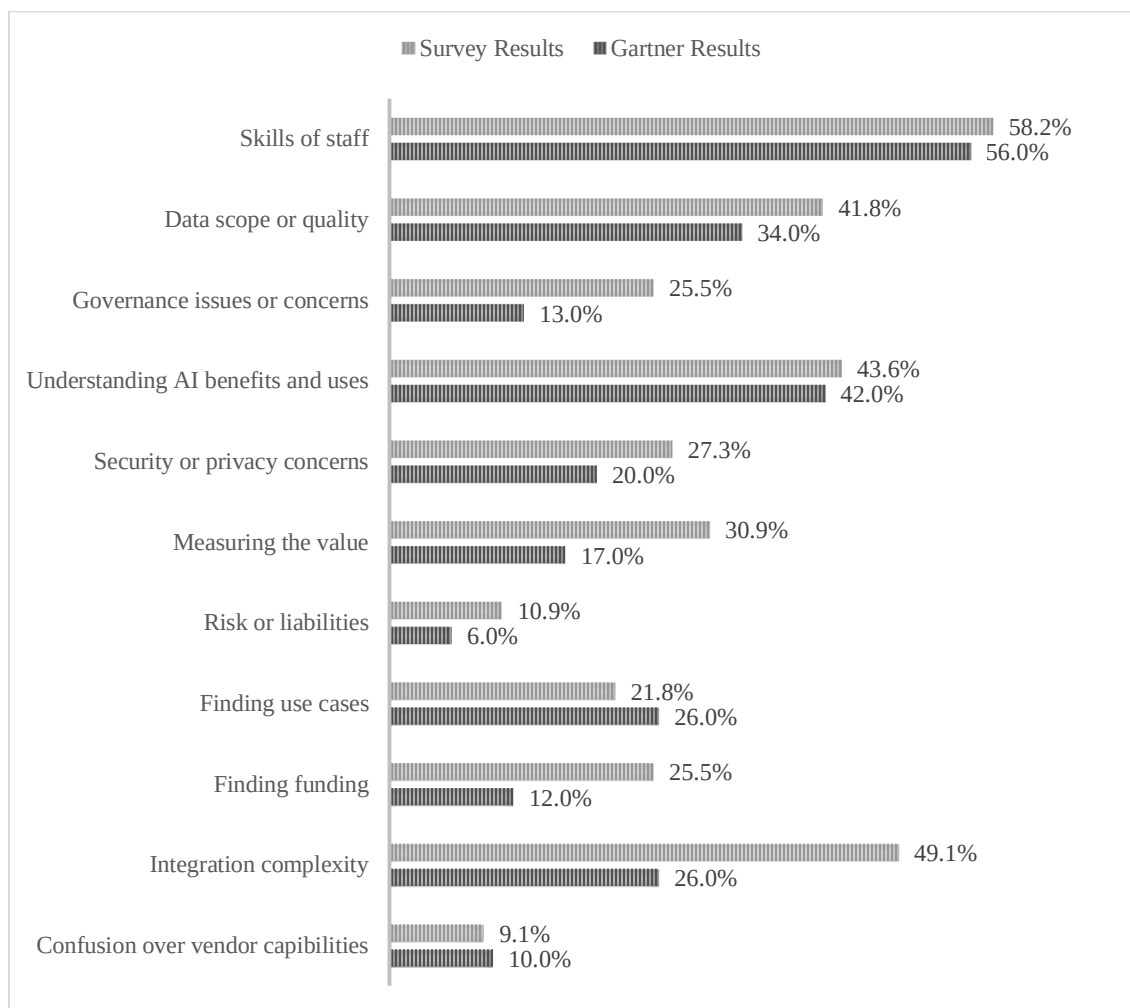


Figure 17: Comparison of Adoption Barriers to Machine Learning in Enterprise Survey Results in Pakistan - 2020 with Gartner Survey Results - 2019 [9]

Fig. 17 shows the results for adoption barriers of Machine Learning in enterprises that participated in our survey in 2020 and the survey conducted by Gartner in 2019 [9]. In both these survey, the 3 common adoption barriers that are amongst the top 5 barriers to adoption of Machine Learning are as follows:

1. Skills of Staff
2. Understanding AI benefits and uses
3. Data cope or quality

Integration complexity is found to be a much bigger challenge in Pakistan's survey results compared to Gartner results in [9]. This can be because the current state of digitalization in

Pakistan is not comparable to the rest of the world. Most of the systems are working in silos and require considerable efforts and effective strategy to be integrated.

Besides this, governance issues, measuring the value Machine Learning add to the business and finding funding are also more serious issues in Pakistan compared to rest of the world.

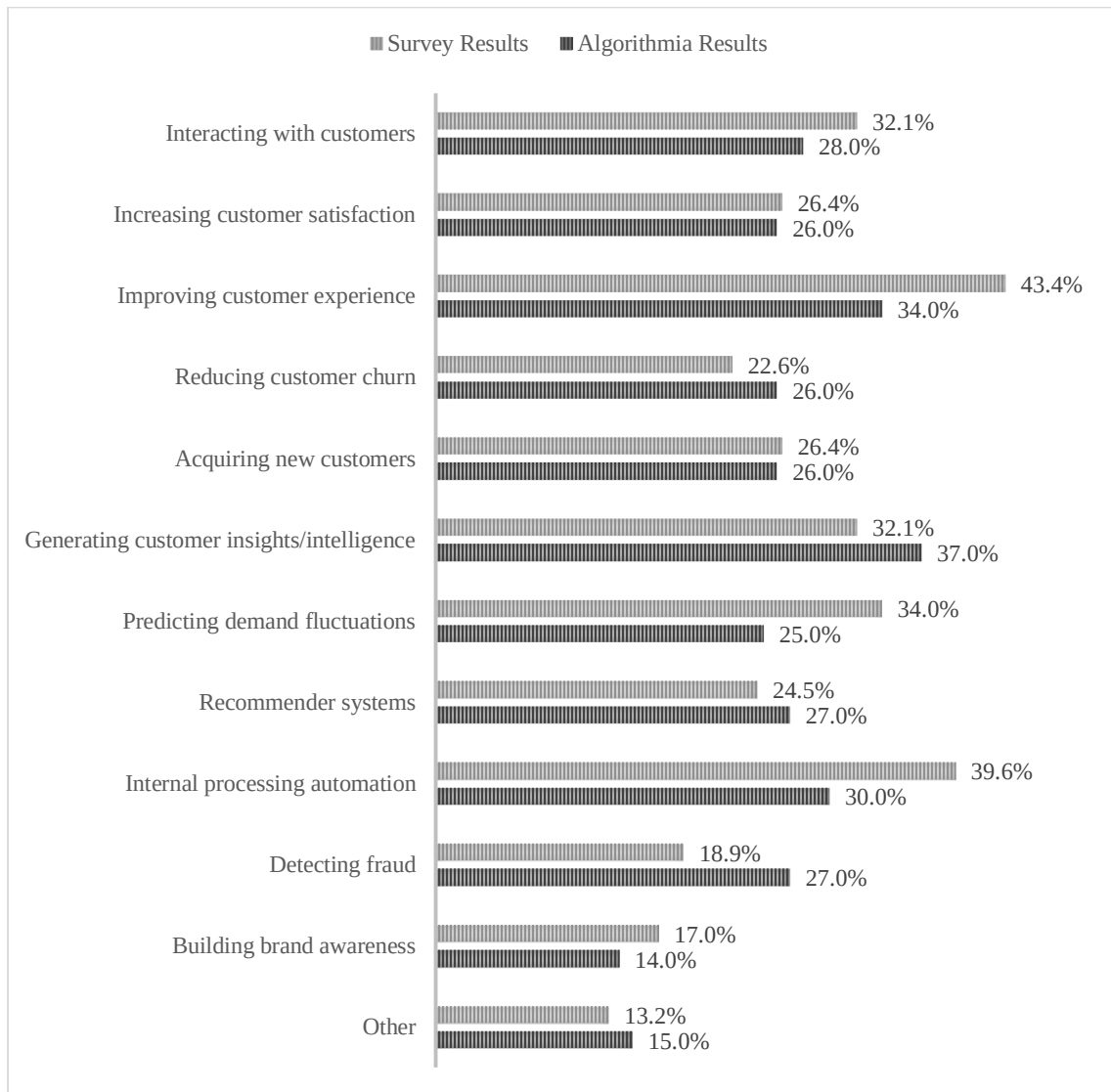


Figure 18: Comparison of Machine Learning Use cases in Enterprise Survey Results in Pakistan - 2020 with Algorithmia Survey Results - 2020 [7]

Fig. 18 shows the results for use cases of Machine Learning in enterprises that participated in our survey and the survey conducted by Algorithmia [7]. The percentage shares of use cases in both the surveys are quite similar.

Improving customer experience, predicting demand fluctuations and internal process automation have much percentage shares in use cases of Pakistan compared to those in Algorithmia results [7]. On the other hand, fraud detection has a much smaller percentage in Pakistan's results i.e. 18.9% compared to its share in Algorithmia results i.e. 27.0% [7]. This shows that Pakistan's is yet to tap the advantages of using Machine Learning for fraud detection.

7. Discussion

The survey results indicate that the state of Machine Learning in Enterprises in Pakistan is much more in early development stages compared to its state around the globe.

Barriers to Adoption of Machine Learning by Organizational Maturity

The maturity-wise results of barriers to adoption shown in Fig. 19 below suggest that there are much more challenges for enterprises in early stages compared to more mature ones in Pakistan. Skills of staff, security or privacy concerns and having not sufficient funding are major problem faced by enterprises in their earlier stages. Whereas, integration complexity and scope or quality of available data are major concerns for more mature organizations.

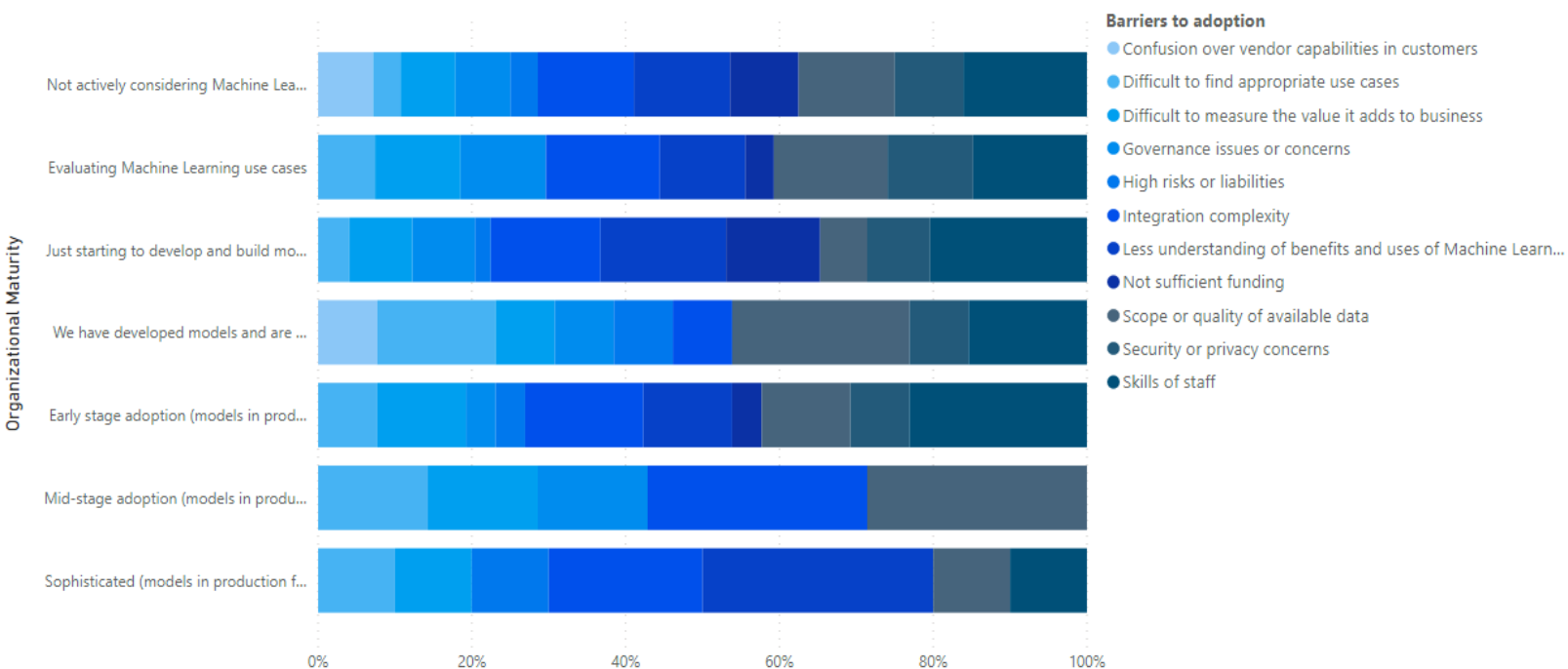


Figure 19: Survey Results – Organizational Maturity-wise Adoption Barriers to Machine Learning in Enterprises in Pakistan, 2020

Besides this, following concerns about Machine Learning are common among all the maturity levels of enterprises in Pakistan:

1. Difficult to find its appropriate use cases
2. Difficult to measure the value it adds to the business
3. Less understanding of its benefits and uses

Therefore, Educational institutions can play a major role in addressing these concerns.

Use Cases of Machine Learning by Organizational Maturity

A good indication for Pakistan is that there are some common use cases which are being considered across all maturity levels of enterprises as shown in Fig. 20 below. These include acquiring new customers, improving customer experience, generating customer insights, predicting demand fluctuations, recommender systems and reducing customer churn. This shows that use cases of Machine Learning in enterprises in Pakistan are more customer-centric irrespective of the maturity level of the organization.

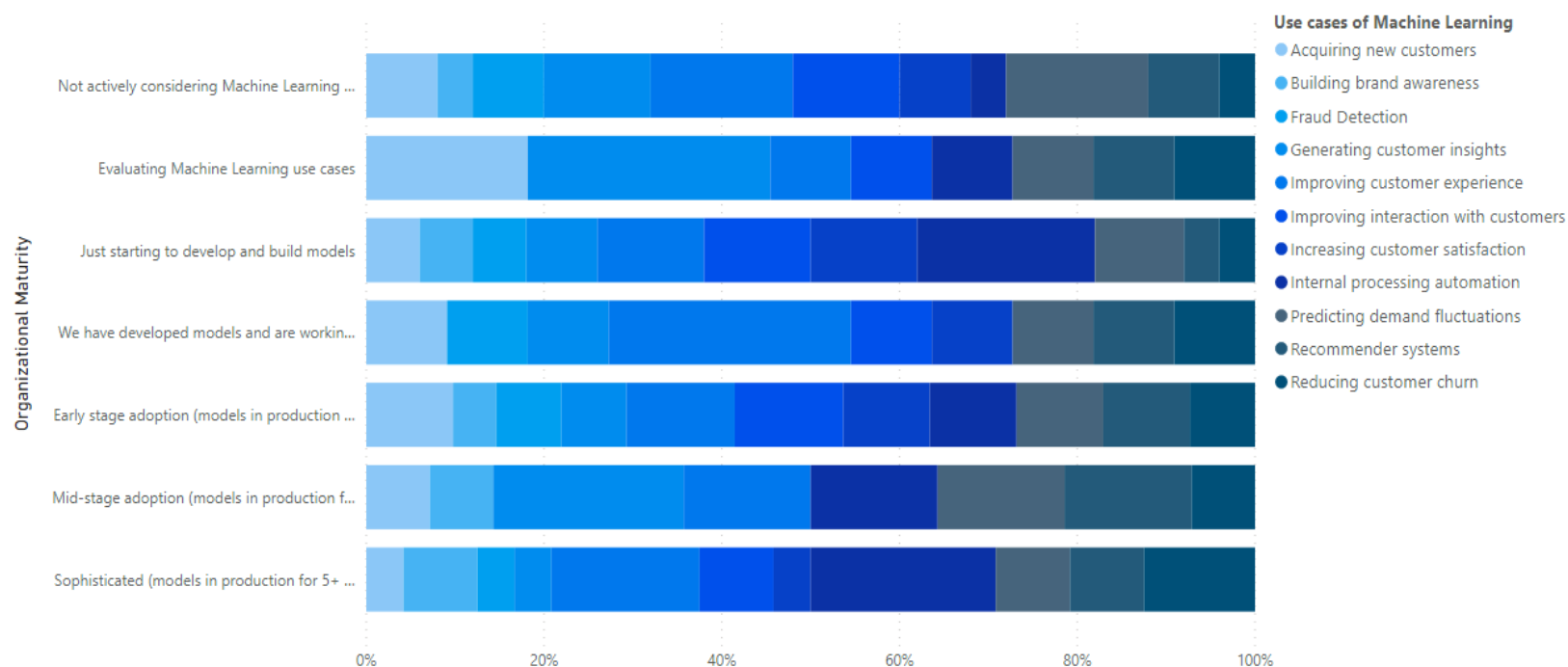


Figure 20: Survey Results – Organizational Maturity-wise Use Cases of Machine Learning in Enterprises in Pakistan, 2020

Fraud detection and increasing customer satisfaction are majorly being focused on by early-stage enterprises. Whereas, using Machine Learning to build brand awareness and internal process automation is more common in mid-stage to sophisticated enterprises

Machine Learning Techniques by Organizational Maturity

Majority of the organization irrespective of their maturity level use Supervised or Semi-supervised learning techniques for Machine Learning. A good reason for this can be lack of resources required for Unsupervised and Reinforcement learning. However, it can be observed in Fig. 21 below that early to mid-stage enterprises are experimenting more with Un-supervised and Reinforcement Learning techniques compared to sophisticated enterprises. This suggests that use of these learning techniques is growing and they are going to be more popular in sophisticated enterprises in future.

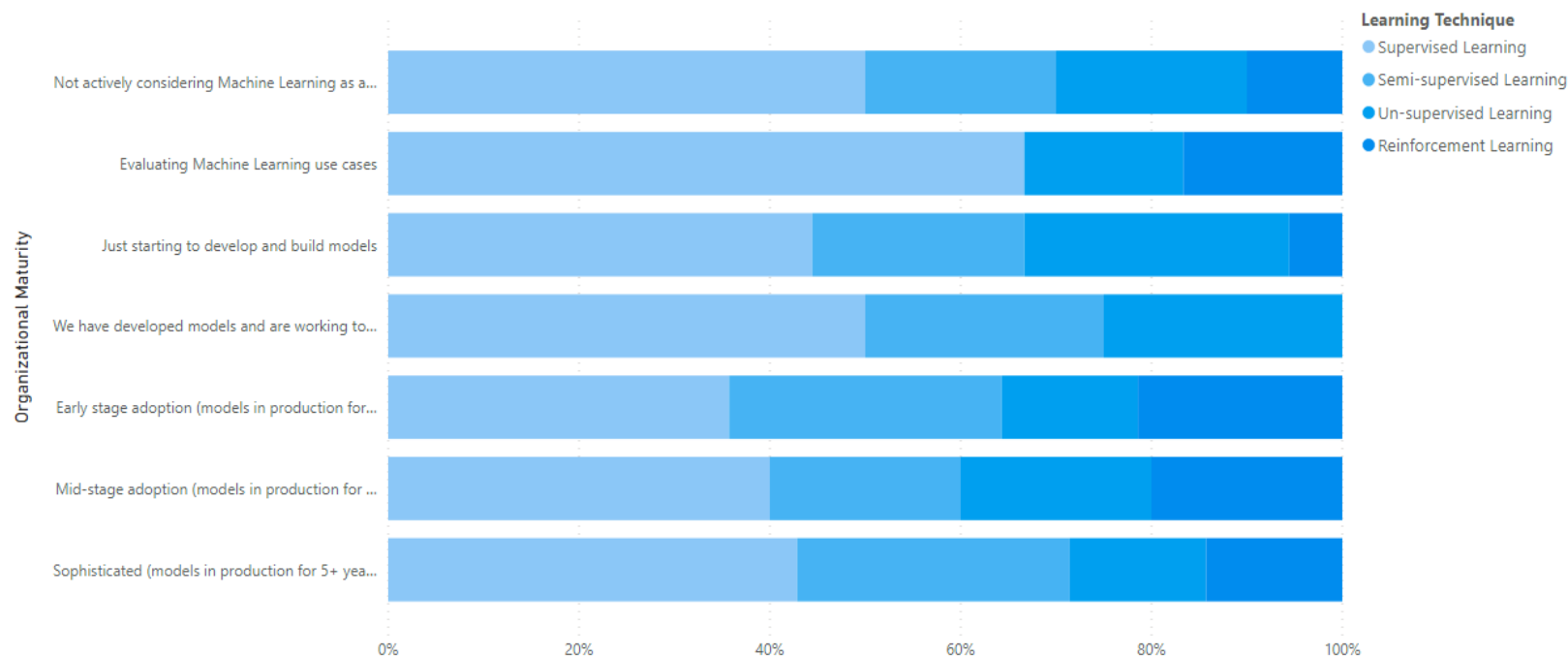


Figure 21: Survey Results – Organizational Maturity-wise Machine Learning Techniques used in Enterprises in Pakistan, 2020

Human Resource Allocation (%), Budget Allocation (%) and Customer Demand (%) for Machine Learning by Organizational Maturity

Human Resource Allocation (%) is mostly 0 – 10 % across all maturity levels of enterprises in Pakistan as shown in Fig. 22 below with very high concentrations in early to mid-stage level enterprises. 60% of even the sophisticated enterprises have only 20 – 30 % human resource allocated for Machine Learning. Only about 20% of the sophisticated enterprises have 90 – 100 % human resource dedicated for Machine Learning.

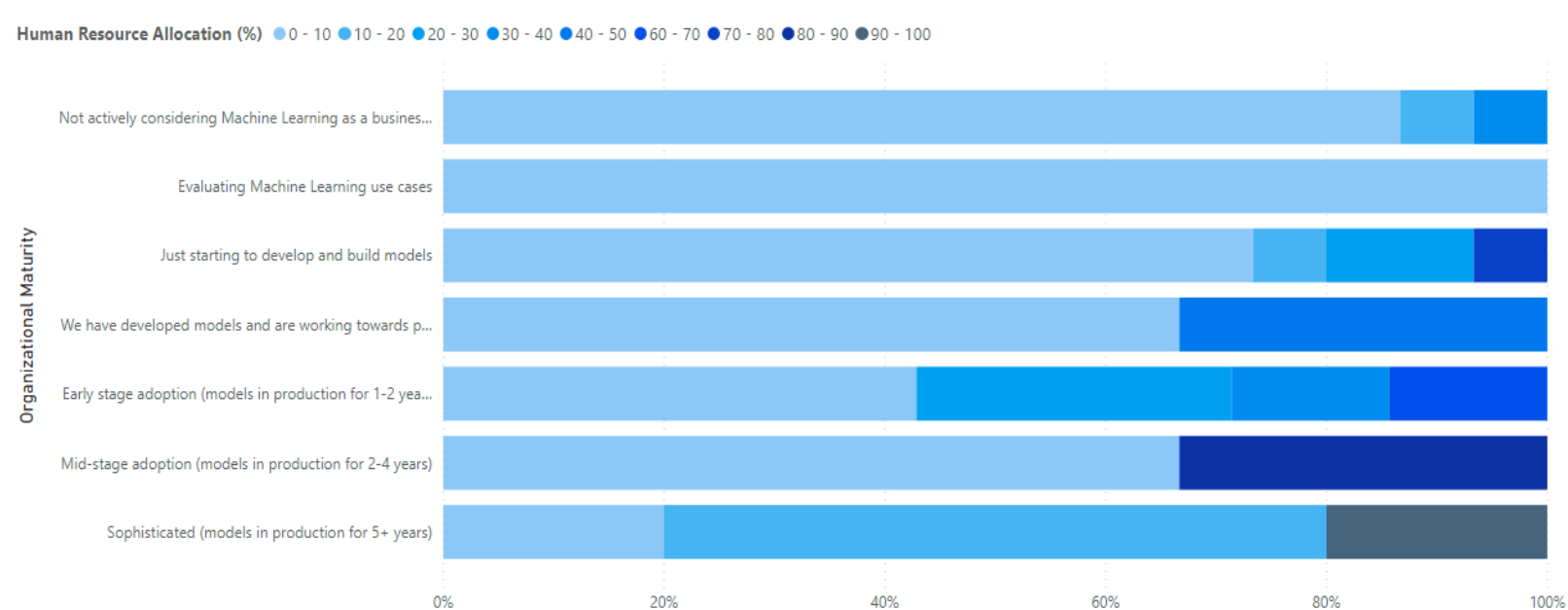


Figure 22: Survey Results – Organizational Maturity-wise Human Resource Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020

These observations are supported by the Budget Allocation (%) and Customer Demand (%) for Machine Learning in Enterprises in Pakistan shown in Figs. 23 and 24. This means that as the demand for Machine Learning will increase more and more companies will grow in their budget and human resource allocations for it as there seems to be a positive correlation between them.

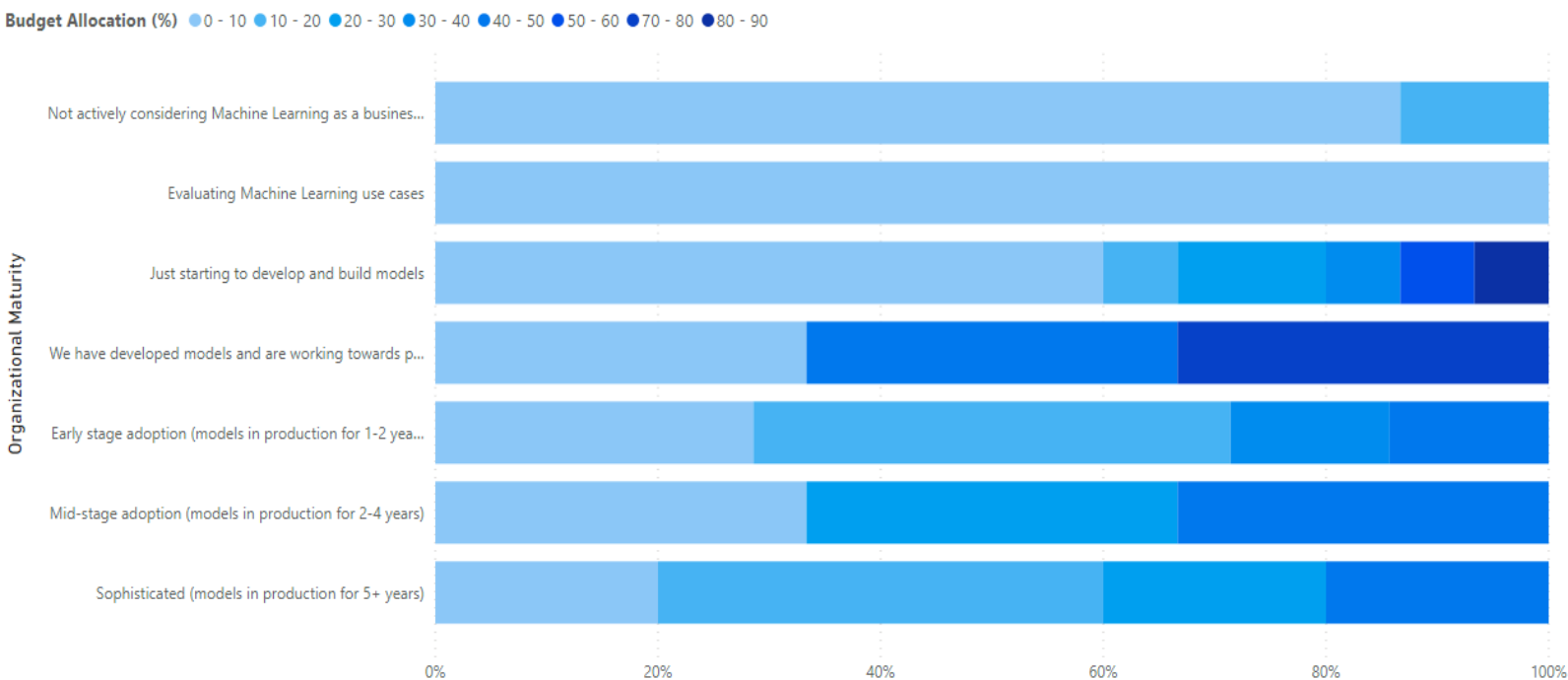


Figure 23: Survey Results – Organizational Maturity-wise Budget Allocation (%) for Machine Learning in Enterprises in Pakistan, 2020

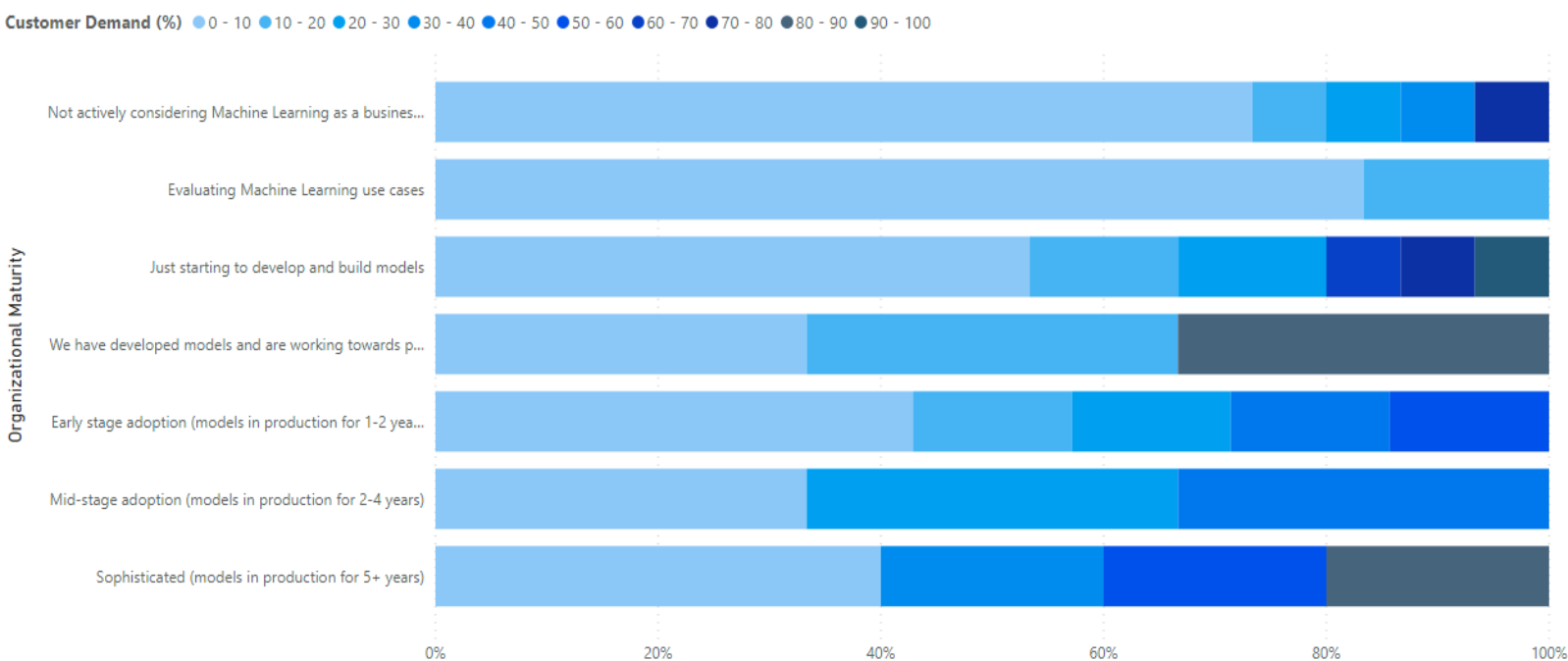


Figure 24: Survey Results – Organizational Maturity-wise Customer Demand (%) for Machine Learning in Enterprises in Pakistan, 2020

Conclusion

In this paper, results have been presented of a survey which was carried out with 300 enterprises of Pakistan in 2020. The questions asked in the survey focused on maturity level of enterprises in Machine Learning, adoption barriers to Machine Learning in these enterprises and use cases of Machine Learning in them. Information like budget allocation, human resource allocation and customer demand of Machine Learning in these enterprises was also collected to get a better understanding. The results show that Machine Learning in Pakistan is still in its early stage of development. However, with the use cases and adoption barriers that have been identified by the enterprises in Pakistan and the growing number of Machine Learning developers being produced by Educational institutions, the authors believe that Machine Learning is going to grow significantly in Pakistan in forthcoming years. The authors also understand that the generalizability of these results is limited because the number of participants in this survey is small. Therefore it is suggested that further surveys are conducted in forthcoming years to establish if there are any patterns in this state of Machine Learning.

References

- [1] R. N. Thomas and R. Gupta, "A Survey on Machine Learning Approaches and Its Techniques;," in *2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, Bhopal, India, 2020.
- [2] J. Brownlee, "A Tour of Machine Learning Algorithms by Jason Brownlee," Machine Learning Mastery, 2019. [Online]. Available: <https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms>. [Accessed 6 November 2020].
- [3] GitHub, "The 2020 State of the OCTO-VERSE," 2020. [Online]. Available: <https://octoverse.github.com/>. [Accessed 13 January 2021].
- [4] P. Vats and K. Samdani, "Study on Machine Learning Techniques In Financial Markets," in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, Pondicherry, India, 2019.
- [5] Muazma, "Pakistan's Rising Data Science Community," kaggle, 2019. [Online]. Available: <https://www.kaggle.com/muazmaz/pakistan-s-rising-data-science-community>. [Accessed 11 January 2021].
- [6] kaggle, "Kaggle's State of Data Science and Machine Learning 2019," 2019. [Online]. Available: <https://www.docdroid.com/qzyxCr4/kaggle-state-of-data-science-and-machine-learning-2019-pdf>. [Accessed 11 January 2021].
- [7] ALGORITHMIA, "2020 State of Enterprise Machine Learning," 2020. [Online]. Available: https://info.algorithmia.com/hubfs/2019/Whitepapers/The-State-of-Enterprise-ML-2020/Algorithmia_2020_State_of_Enterprise_ML.pdf. [Accessed November 2020].
- [8] D. Curry, "Gartner: Future Of AI, And The Challenges," RTInsights.com, 14 November 2019. [Online]. Available: <https://www.rtinsights.com/gartner-future-of-ai-and-the-challenges/>. [Accessed 10 January 2021].
- [9] Gartner, "3 Barriers to AI Adoption," Smarter With Gartner, 18 September 2019. [Online]. Available: <https://www.gartner.com/smarterwithgartner/3-barriers-to-ai-adoption/>. [Accessed 10 October 2020].
- [10] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255 - 260, 2015.
- [11] K. Wiggers, "Gartner: Enterprise use of AI grew 270% over the past 4 years," The Machine, 21 January 2019. [Online]. Available: <https://venturebeat.com/2019/01/21/gartner->

- enterprise-ai-implementation-grew-270-over-the-past-four-years/. [Accessed 23 December 2020].
- [12] Gartner, "Survey Analysis: AI and ML Development Strategies, Motivators and Adoption Challenges," 19 June 2019. [Online]. Available: <https://www.gartner.com/en/documents/3940005/survey-analysis-ai-and-ml-development-strategies-motivat>. [Accessed 16 January 2021].
- [13] ALGORITHMIA, "Machine learning use cases," Algorithmia, 6 December 2019. [Online]. Available: <https://algorithmia.com/blog/machine-learning-use-cases>. [Accessed 20 January 2021].
- [14] M. I. Jordan, "Machine learning: trends, perspectives and challenges," in *TURC '18: Proceedings of ACM Turing Celebration Conference*, China, 2018.
- [15] S. Ray, "A Quick Review of Machine Learning Algorithms," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Faridabad, India, 2019.
- [16] O. Obulesu, M. Mahendra and M. ThirilokReddy, "Machine Learning Techniques and Tools: A Survey," in *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2018.
- [17] M. Mohammed and M. B. Khan, *Machine Learning Algorithms and Applications*, CRC press Taylor and Francis Group, 2017.
- [18] S. R. Guruvayur and R. Suchithra, "A detailed study on machine learning techniques for data mining," in *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, India, 2017.
- [19] Zhou, L. &. Pan, S. &. Wang, J. &. Vasilkos and Athanasios, "Machine Learning on Big Data: Opportunities and Challenges," *Neurocomputing*, 2017.
- [20] A. L'Heureux, K. Grolinger, H. F. E. Yamany and M. A. M. Capretz, "Machine Learning With Big Data: Challenges and Approaches," *IEEE Access*, 2017.

Appendix – Research Survey

Personal Information

- Full Name
- Organization Name
- Designation
- Email ID

Use Cases of Machine Learning

Q. What are the use cases of Machine Learning that your organization has been working on?

- ☐ Improving interaction with customers
- ☐ Increasing customer satisfaction
- ☐ Improving customer experience
- ☐ Reducing customer churn
- ☐ Acquiring new customers
- ☐ Generating customer insights
- ☐ Recommender systems
- ☐ Predicting demand fluctuations
- ☐ Internal processing automation
- ☐ Building brand awareness
- ☐ Fraud Detection
- ☐ Reducing Costs
- ☐ Other

Enterprise Maturity Level

Q. How mature is your organization in the field of Machine Learning?

- Not actively considering Machine Learning as a business solution
- Evaluating Machine Learning use cases
- Just starting to develop and build models
- We have developed models and are working towards production
- Early stage adoption (models in production for 1-2 years)
- Mid-stage adoption (models in production for 2-4 years)
- Sophisticated (models in production for 5+ years)

Human Resource Allocation (%)

Q. What percentage of your organization's human resource is allocated for Machine Learning solutions?

- | | |
|-----------|------------|
| ○ 0 – 10 | ○ 50 – 60 |
| ○ 10 – 20 | ○ 60 – 70 |
| ○ 20 – 30 | ○ 70 – 80 |
| ○ 30 – 40 | ○ 80 – 90 |
| ○ 40 – 50 | ○ 90 – 100 |

Budget Allocation (%)

Q. What percentage of your organization's budget is allocated for Machine Learning solutions?

- | | |
|-------------------------------|--------------------------------|
| ☐ 0 – 10 | ☐ 50 – 60 |
| ☐ 10 – 20 | ☐ 60 – 70 |
| ☐ 20 – 30 | ☐ 70 – 80 |
| ☐ 30 – 40 | ☐ 80 – 90 |
| ☐ 40 – 50 | ☐ 90 – 100 |

Customer Demand (%)

Q. What percentage of your organization's customers demand Machine Learning solutions?

- | | |
|-------------------------------|--------------------------------|
| ☐ 0 – 10 | ☐ 50 – 60 |
| ☐ 10 – 20 | ☐ 60 – 70 |
| ☐ 20 – 30 | ☐ 70 – 80 |
| ☐ 30 – 40 | ☐ 80 – 90 |
| ☐ 40 – 50 | ☐ 90 – 100 |

Adoption Barriers

Q. What are the barriers to adoption of Machine Learning in business solutions in your organization?

- ☐ Skills of staff
- ☐ Scope or quality of available data
- ☐ Integration complexity
- ☐ Governance issues or concerns
- ☐ Security or privacy concerns
- ☐ Less understanding of benefits and uses of Machine Learning
- ☐ Confusion over vendor capabilities in customers
- ☐ High risks or liabilities
- ☐ Difficult to measure the value it adds to business
- ☐ Difficult to find appropriate use cases
- ☐ Not sufficient funding
- ☐ Other

Machine Learning Techniques Used

Q. Which of the following groups of Machine Learning Techniques have been used in your organization?

- ☐ Supervised Learning
- ☐ Semi-supervised Learning
- ☐ Reinforcement Learning
- ☐ Un-supervised Learning

Machine Learning Algorithms Used (Based on Similarity)

Q. Which of the following groups of Machine Learning Algorithms have been used in your organization?

- ☐ Regression Algorithms
- ☐ Instance-based Algorithms
- ☐ Regularization Algorithms
- ☐ Decision Tree Algorithms
- ☐ Bayesian Algorithms
- ☐ Clustering Algorithms
- ☐ Association Rule Learning Algorithms
- ☐ Artificial Neural Network Algorithms
- ☐ Deep Learning Algorithms
- ☐ Other