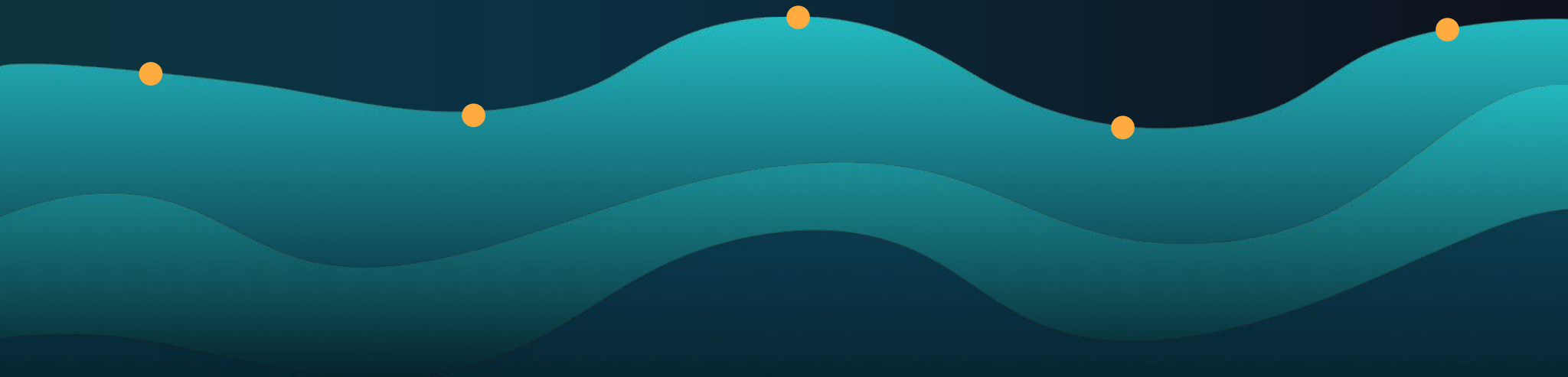


Q1 2022 MLR Report



By Chen Song,
Data Scientist, Loxz Digital

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



Preface

Remaining competitive in a post-pandemic dynamic ML marketplace is one of the biggest challenges that many organizations face. Chances are if you are not using ML at your organization for decision-making in some capacity, the chances of becoming obsolete increase exponentially.

Machine learning (ML), as a core subset of artificial intelligence(AI), helps organizations enhance business agility, value and scalability. According to McKinsey, "AI could potentially deliver an additional economic output of around \$13 trillion by 2030, boosting global GDP by about 1.2% a year." (Young, S., 2021, March 11). An organization's ability to adopt machine learning and evaluate machine learning readiness is crucial to determining market positioning, investments, and business value to remain competitive because the impact and insights garnered from machine learning illuminates business processes across all industries. Companies are increasing the pace of adoption in machine learning, realizing the business value of machine learning in the marketplace. According to the report Top Machine Learning Trends for 2022, the value of the machine learning market will increase to \$117 billion by 2027 from \$8 billion in the year 2021. (Kapoor, A., 2022, February 8).

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report

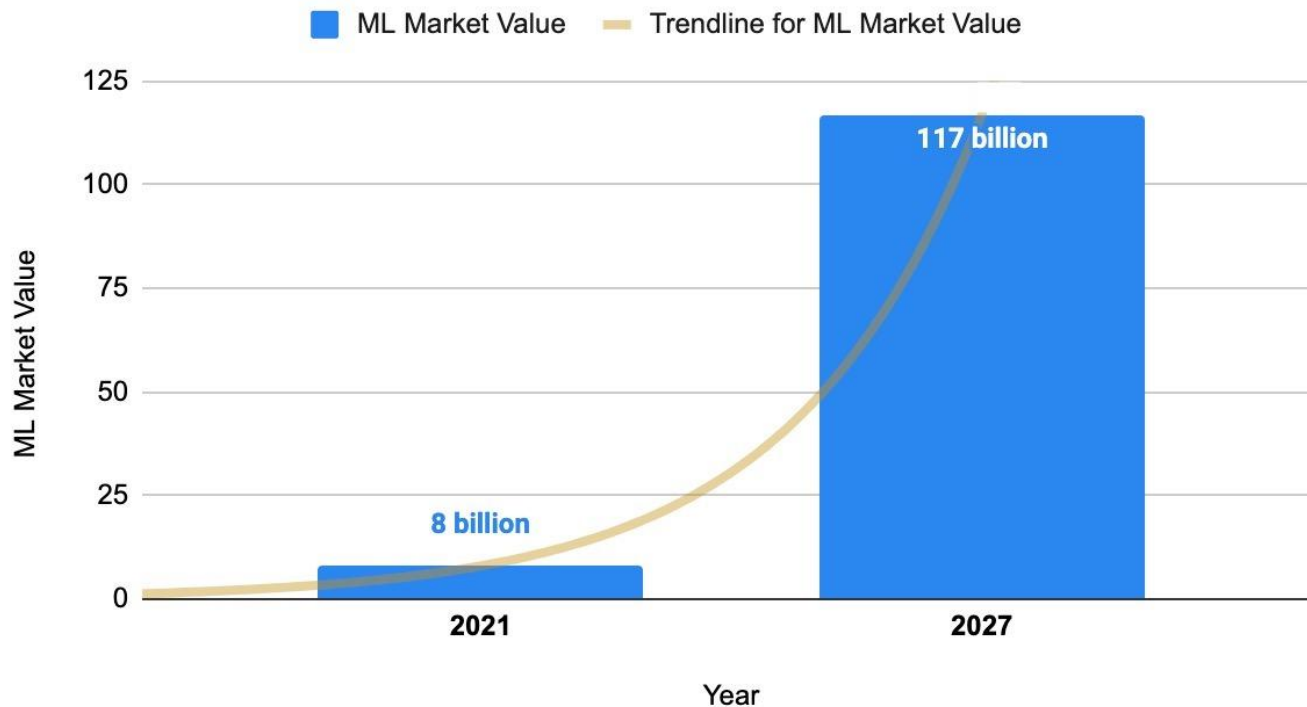


Figure 1. ML market value increase from 2021 to 2027

Loxz Digital continues to help organizations, academic institutions, and individuals buttress ML development, and spot potential weaknesses or opportunities by providing real-time diagnostic scores and insights from a machine learning lifecycle market perspective.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



By taking [Loxz Digital's machine learning readiness survey-organization version](#), you can get a scientific diagnostic evaluation regarding the machine learning readiness of your organization. We provide you with a machine learning readiness (MLR) overall score and granular sub-scores across the machine learning lifecycle (See [Table 1](#) and [Graphic 1](#)). With the MLR score and subscores, diagnostic insights allow organizations view in granularity the advantages and disadvantages of the current model development cycle regarding machine learning readiness. From quarter two of 2022, diagnostic insights will also available on Loxz Digital's MLR dashboard. There is a beta version prepared for both Organizations and academic institutions.

Loxz Digital has also designed and developed an academic version of the machine learning readiness survey. We are pleased to announce that the [machine learning readiness survey-student version](#) is available now: Email Mina Mahdavi, PhD, to get your academic institution or organization on the list and we will create a separate subdomain for you. Mina can be reached at Mina@loxz.com

As part of our current pilot programs, [the University of California, Irvine](#), and [Sweet Briar College](#) show great interest in the student machine learning readiness survey. By customizing the machine learning readiness survey for students, Loxz Digital provides a practical and diagnostic tool for aspiring ML data scientists to self-evaluate their machine learning readiness and for both academic institutions and the HR Department to evaluate the curriculum and aptitude of candidates for greater efficiencies. We believe that by hosting the student machine learning readiness survey on their websites and analyzing the MLR score, subscores, and insights, students can understand and evaluate their knowledge and experiences in machine learning, and academic institutions can evaluate the current curriculum correspondingly.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



Student MLR Survey

[The Loxz Digital Machine Learning Readiness \(MLR\)-Student version](#) survey measures the student's knowledge and experience in the field of machine learning. The items within this assessment act as a proxy for success in machine learning and data science careers. The survey consists of 6 sections, including general questions, data preparedness questions, modeling questions, career trajectory questions, ML aptitude questions, and business value questions. The 6 sections allow us to perform in-depth analysis and diagnosis of machine learning readiness from different perspectives. Similar to the organization version of the machine learning readiness survey, the student survey also provides an overall score and innovative subscores. The value is distinguished in the sub-scoring methodology.

In the student survey, we introduced sets of open-ended questions. (We created our own set of answers for each open-ended question called "Answer Corpus". The goal here is to measure the semantic similarity between the survey input answer and our Answer Corpus. The approach is to use transfer learning with the latest pre-trained BERT tokenizer and a BERT model. We first tokenize the student answer and then each answer in the Answer Corpus then feed those tokens into the BERT model. Then we capture the dense vector embeddings from the last hidden state. To get the result, we mean pool the masked embeddings to get the vector representations of answers for calculating the cosine similarity score between the student answer and answers in Answer Corpus. The final similarity score will be averaged among all student-corpus answer pairs and normalized to 0 to 100.)

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



The student version of the machine learning readiness survey was designed and developed to measure MLR at different levels of granularity, aiming to provide an assessment tool for individual students, academic institutions and HR departments.

Individual student level.

- The student version of the machine learning readiness survey measures a student's ML aptitude and the ability to design, construct, and monitor machine learning models that capture a business value. The MLR score and subscores allow the students who took the survey to get an overview of their knowledge and experiences in the machine learning field. The insights coming along with the scores provide the students with diagnostic details about their advantages and disadvantages over competitors.

Academic institutions level.

- The student version of the machine learning readiness survey also provides a tool for academic institutions because it measures a school's ability to offer a data science/ Machine learning program that is designed based on industry needs. On one hand, academic institutions are able to adjust course design based on the survey results, using the MLR score, subscores, and insights as a reference. On the other hand, getting to know students' knowledge and experience in the field of machine learning is also an effective way for schools to get feedback from students regarding their degrees of digesting course contents.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



HR Department level

- The student version of the machine learning readiness survey provides the HR department with an assessment tool when interviewing new graduates candidates in machine-learning-related roles. It also provides a QR code that HR departments can parse automatically and use for resume segmentation, relating to ML sub-categorical scoring. It will parse all five categories of ML lifecycle, and a career trajectory score that should help HR teams evolve and be more efficient.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATION
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



MLR Pearson Correlations Heatmap

It's a pleasure to announce that our new [Machine learning readiness survey dashboard](#) is available from Quarter 1, 2022. New elements were added to the machine learning readiness survey dashboard to present insights and information retrieved from the survey responses. As a key component of Loxz Digital's MLR survey dashboard, our MLR Pearson correlation Heatmap reveals the scoring mechanism from the question relations perspective.

- What is MLR Pearson Correlations Heatmap

MLR Pearson correlations heatmap allows the audiences to get a deeper understanding of how the MLR score are correlated in each question. With the MLR Pearson correlations heatmap, respondents, do not blindly go through the entire survey and get an MLR score anymore. Instead, they know about how their answers contribute to the MLR score. Our goal is to provide insights on how those question features are correlated for a certain group of survey takers. The correlation matrix is computed by calculating the Pearson Coefficients between any two features in the filtered survey data. The data are filtered in two ways: filtering by user group or filtering by question group (subcategories). For more information and methodology about filtering and Correlation Matrix, Please read [MLR Pearson Correlations Heatmap Methodology](#). This paper describes in detail the mathematical and statistical method of the MLR Pearson correlations heatmap.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
MLR Report


loxz digital

- Preview of MLR Pearson Correlations Heatmap

We collected a user-question matrix M , where each entry M_{ij} represents the score that user i got by answering question j . If the correlation value at entry M_{ij} of the heatmap is positive (a warmer hue) then it means there is a positive linear correlation between the scoring of question i and question j . On the other side, a negative value (a colder hue) means there is a negative correlation between the scoring of question i and question j . The larger the magnitude of the value is, the stronger the correlation is, and no correlation with the value being 0.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report

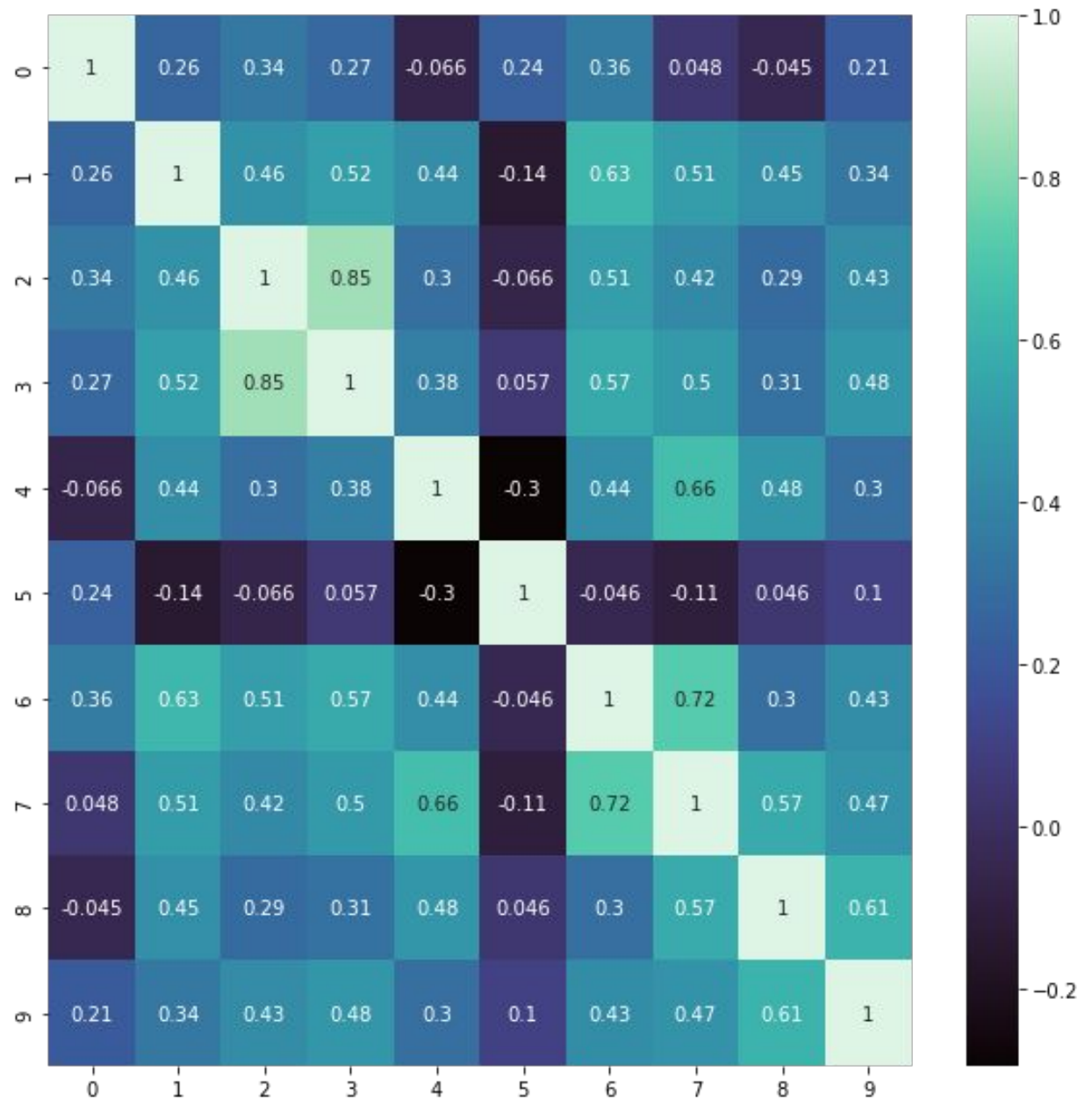


Figure 2. MLR Pearson correlation Heatmap

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

MLR Key Insights of Quarter 1 2022

- The prospect of development in deep learning brings about diversification of data types that organizations are capable of analyzing.

Enterprises capable of deep learning become more intelligent and diversified as they wrangle data from different industries. In part due to the development of deep learning technologies, including Convolutional Neural Networks(CNN), Vision Transformers, NLP models such as BERT, XLNet, Roberta, etc. The development of these technologies enables organizations to apply machine learning to a broader domain scope because they are capable of analyzing more complex data under more sophisticated scenarios. The results of our survey (Figure 2) show that the data the organizations are collaborating with have various types. The categorical data, (including image, video, and audio data), and text data are the most popular and common data types that organizations are engaging with. With the ML cultural shift to real-time ML, time-series data and time-series machine learning models have gained in popularity, and time-series analysis is becoming a handy tool for the forecasting model because of its ability to filter out the noises, learn the pattern and make predictions in one set. However, due to the “historical-data-based” nature of time-series analysis, time-series analysis is also a great challenge for data scientists. It also presents the greatest model monitoring challenges.

Q1 2022
ML Readiness Survey Report

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF Q1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



The non-stationary nature, together with the “historical-data-based” nature of the time series data and analysis makes it more challenging to monitor the time-series models. On one hand, time series predictions are based on the assumption that the future trend of the event will stay similar to the historical trend, and this is the “historical-date-based” nature. It requires the integration of judgment and statistical methods when developing and monitoring the model because too strict assumptions will lead to an unreliable forecast, and too flexible assumptions will lead to uncertainties in the predictions. Either unreliable models or uncertain predictions put more challenges in the monitoring process. On the other hand, some static statistical data are important metrics in model monitoring, such as mean value, median value, etc. However, in the time series model, these statistical properties change over time. The result from our survey shows ~10% less in time-series data type compared to text data type and one potential reason is because of the challenges of time-series analysis.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report

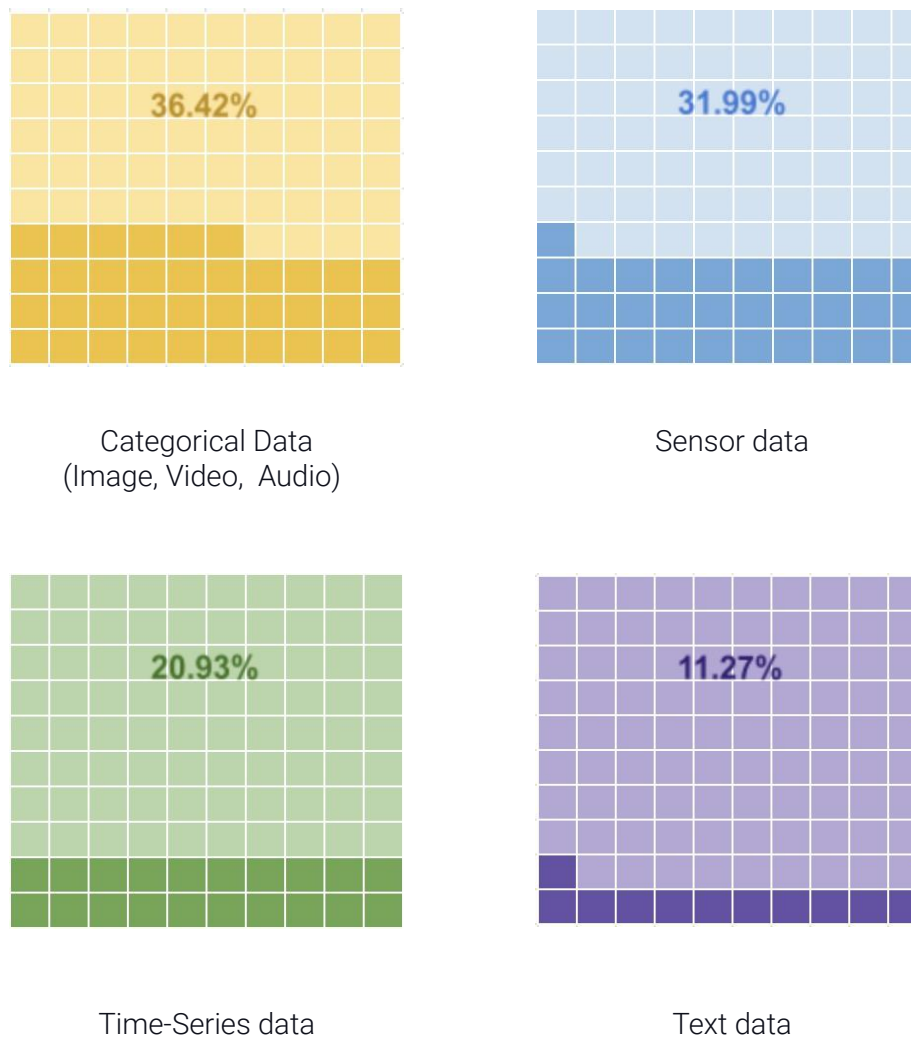


Figure 3. Data types proportions across survey takers

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



- Model deployment is one of the most important and most challenging processes in the machine learning lifecycle. Leaders and innovators are confident about their deployment results, ~75% of leaders and innovators indicate that the deployment result met or exceeded their original expectations. However, ML performers suffer from uncertainties. Any organizations that have adopted machine learning in their businesses can not attach too much significance to the model deployment, because model deployment allows organizations to actually integrate ML models into practical decision-making processes. However, deploying models also means challenges. To gain the practical power of machine learning models, organizations need to deploy their models into products, and the deploying process not only requires financial support but it also requires consistent model monitoring and multi-departmental collaboration, involving data scientist teams, IT teams, software developers, and business teams. Model deployment is closely associated with monitoring as well. ML leaders usually have stable business models, capital flow, and data streaming. Moreover, their mature functional teams are able to ensure the effectiveness of the deployment process. So as shown in Figure 3, 75% of the leaders show that the deployment results met or even exceeded their expectations and none of them indicate any of their deployments didn't meet the expectations. Compared to ML leaders, ML performers suffer from uncertainty of the deployment, ~60% of ML performers indicate that either it's too early to tell the deployment results or the results are difficult to tell.

Result of Model Deployment

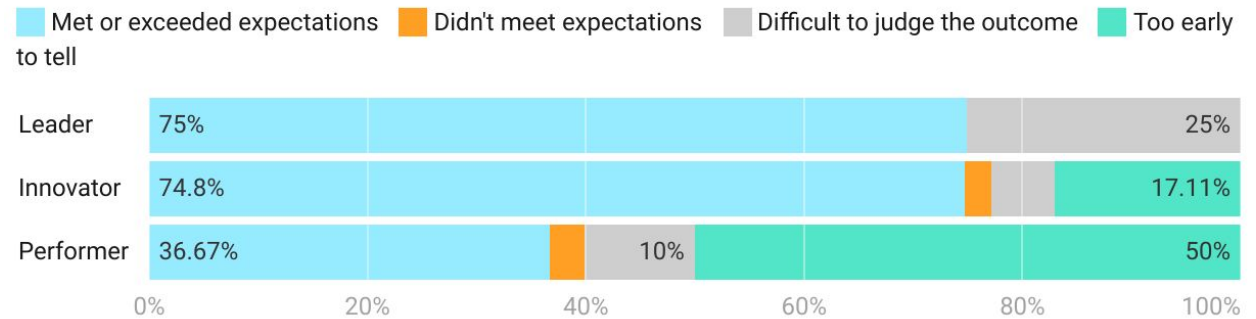


Figure 4. Result of model deployment

- Machine learning serves different organizational needs, from employees to customers, and from technology to marketing. But improving marketing and customer service are the primary goals that enterprises hoping to achieve with ML.

As shown in Figure 4, there are 58.61% of organizations show that the primary goal of adopting ML in their business is to improve marketing and customer satisfaction.

Personalized products and services, eyeball-catching content, and service robots are determining factors to seize the market. The results of our survey illustrate the significant role of machine learning within organizations. Machine learning is also applied to enterprise management and business efficiency largely. There ~20% of the respondents indicated that they are using machine learning primarily to automate routine processes. And there are 16% of the respondents showed they are using machine learning primarily for enterprise internal management.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report

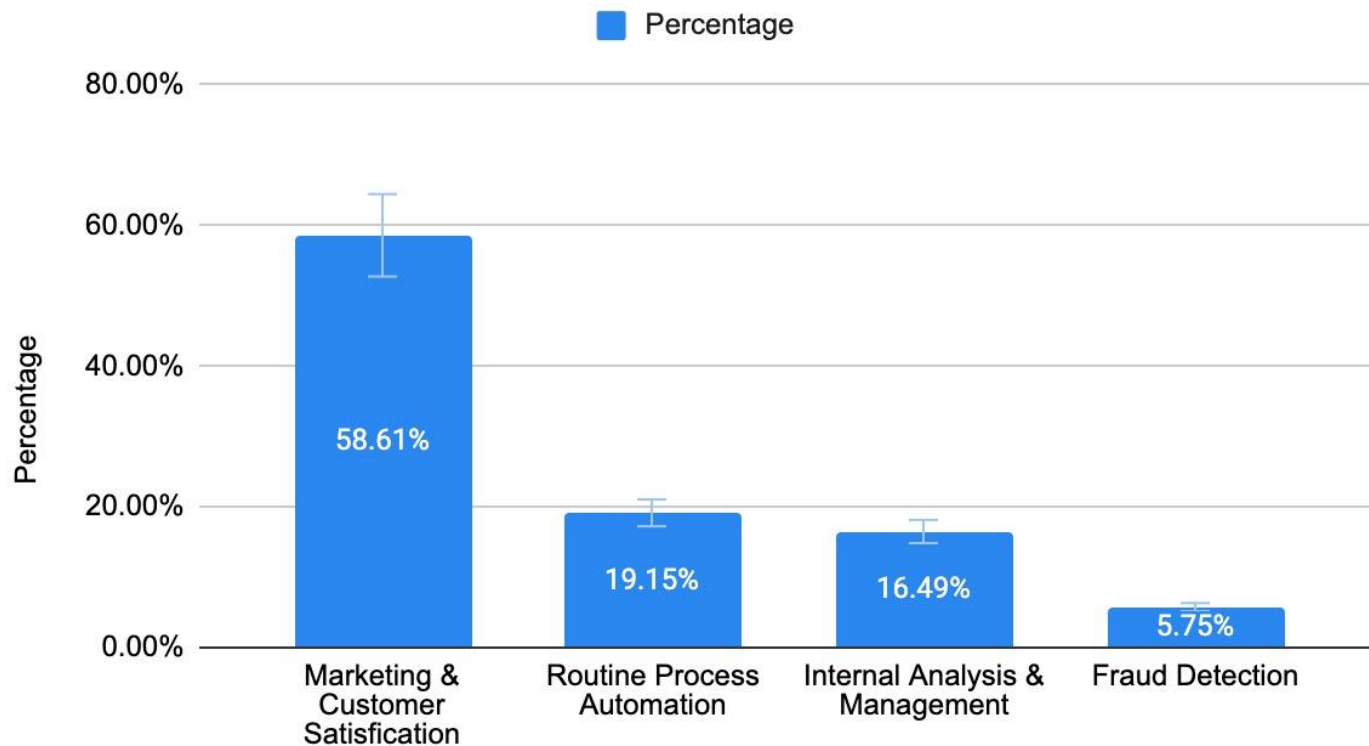


Figure 5. The goal of adopting machine learning

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

- In the era of big data, data plays an influential role in any business development. And data preparation of an organization is scarcely influenced by how important it is of ML to an organization.

As shown in Figure 5, no matter how important is machine learning as a differentiator to an organization (on scale of 2-5, 2 as not important, 5 as extremely important), the average data preparedness scores across all survey takers are at a high level. And the average data preparedness scores do not grow continually or linearly as the differentiator increases.

However, the overall MLR, MD score, DM score, MM score and BV scores all show a linear correlation with the differentiator. The more important machine learning is to an organization, the higher score they got in stages of the ML lifecycle. A whitepaper by Seagate IDC predicts that the Global Datasphere will grow from 33 Zettaby in 2018 to 175 Zettabytes by 2025(IDC White Paper, 2018). As datasphere is growing, organizations may put more emphasis on data preparation.

Q1 2022

ML Readiness Survey Report

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



How important is machine learning (ML) as a differentiator to your organization?



	2	3	4	5
Overall	33	41	48	59
DP	62	56	60	67
MD	23	35	41	53
DM	19	31	42	59
MM	19	49	63	81
BV	56	56	64	70

Note: DP=Data Preparedness; MD=Model Development; DM=Deploying Models; MM=Model Monitoring; BV=Business Value

Figure 6. Machine learning score vs Importance of ML

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



Conclusions

Machine learning is thriving in 2022 with the development of ML tools and techniques. Leah Forkosh Kolben, the Co-founder & CTO at cnvrg.io, believes that Machine learning (ML) models will become easier to develop, implement and maintain, utilizing time-saving tools and turn-key algorithms (Kolben, L., 2022 January). Loxz Digital's Machine learning readiness survey aims to provide you with a tool to efficiently assess your machine learning performance throughout the entire machine learning lifecycle and among competitors. Throughout this report, we highlight our 2 pilot programs of machine learning readiness survey-academic version, and we are providing you key insights into ML adoption trends, the role of machine learning in organizations and the influence of the decade of data for machine learning adoption. We see an obvious trend that machine learning is thriving and many organizations are actively adopting machine learning for both internal use and external use.

Major industries are spawning off of each lifecycle of ML. We're seeing hypertooling in ML Monitoring and a deluge of external datasets being introduced to the industry to buttress model development. At one time, it was more data, now it's the right datasets to introduce. RealTimeML is going to be the showpiece here in the next few quarters, as industries will be thirsty for inferences or predictions served in milliseconds within windows of campaigns. This may or may not marginalize the current ML climate, but realtime and streaming is without a doubt here to stay and is the bedrock of what we do here at Loxz. You are encouraged to check out our [LinkedIn page](#) to get more insights and information on previous quarter reports.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



About us

Loxz Digital Group is a Machine Learning Collective located in Berkeley, CA. Established in December of 2020, our focus is on building and deploying accurate real-time machine learning models with diverse ensemble techniques for enterprise, law enforcement and government entities.

We have partnered with esteemed organizations such as AWS, and TurboSBIR to help us build machine learning models efficiently and coordinate with law enforcement and government entities as a gateway for the commercialization of our RealTimeML predictive inference.

Specifically, RealtimeML is at the bedrock of what we do. Collectively, the current assembled team has over 40 years of ML experience, housing 9 data scientists, and 3 Ph.Ds all located in the United States and Canada. The data acquired from this survey is exclusively first-party data.

©2022 All Rights Reserved. loxz.com

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



Contributors

Chen Song, Data Scientist,
Lead Author, Lead Analyst

Yiming Zhang, Lead Data Scientist,
Pearson Correlation Heatmap Analyst & Author

Yumi Koyanagi, Designer
Report Designer

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



References

1. Young, S. (2021, March 11). Council post: Five reasons why AI/ML should be top of mind in 2021. Forbes. Retrieved April 4, 2022, from <https://www.forbes.com/sites/forbestechcouncil/2021/03/11/five-reasons-why-aiml-should-be-top-of-mind-in-2021/?sh=544331b84556>
2. Kapoor, A. (2022, February 8). Top machine learning trends for 2022. Medium. Retrieved April 8, 2022, from <https://enlear.academy/top-machine-learning-trends-for-2022-6e7071d37130>
3. e7071d37130
4. The digitization of the world from edge to core. (n.d.). Retrieved April 16, 2022, from <https://www.seagate.com/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>
5. Kolben, L. (n.d.). Here's Why Machine Learning will Thrive in 2022. Toolbox.com. Retrieved April 16, 2022, from <https://www.toolbox.com/tech/artificial-intelligence/guest-article/heres-why-machine-learning-will-thrive/>

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report



Table 1. Loxz Digital MLR sub-scores

- Data Preparedness (DP): Quantifies an organization's ability to efficiently and effectively locate, integrate, and leverage business resources to achieve its machine learning objectives.
- Model Development (MD): Measures the frequency and strategy behind how an organization leverages its resources to construct machine learning models to be as accurate as possible.
- Deploying Models (DM): Assesses the infrastructure, scalability, and methodology an organization uses to integrate machine learning models into systems that are in development or already part of their existing technical infrastructure.
- Model Monitoring (MM): Provides a basis for understanding the approach an organization takes in leveraging technical resources to maintain, monitor, and retrain machine learning models that are in production.
- Business Value (BV): Represents the alignment of strategic initiatives and use of machine learning models to enhance one's business.

PREFACE

STUDENT MLR SURVEY

MLR PEARSON CORRELATIONS
HEATMAP

KEY INSIGHTS OF QUARTER 1 2022

RESULTS OF MODEL DEPLOYMENT

CONCLUSIONS

ABOUT US

CONTRIBUTORS

REFERENCES

APPENDIX

Q1 2022
ML Readiness Survey Report

Graphic 1. Average MLR sub-scores across ML roles

