BUSINESS CASE FOR MACHINE LEARNING IN AUTOMOBILTY

Version	Business Draft 1.0
Date	June 1, 2021
Descriptions	First Business Draft for Implementing Machine Learning in Automobility
Authors	Dhamina Sarvathikari, Tixa Hiteshkumar Patel

TABLE OF CONTENTS

DOCUMENT HISTORY 1
EXECUTIVE SUMMARY..3
Issue…………...3
Anticipated Outcomes ....……3
Recommendation 4
Justification………...….4
MARKET ANALYSIS...….4
Market requirements………..5
PROBLEM DEFINITION ....…..6
Problem Statement ....…….6
Organizational Impact..….….6
Technology Migration...………..6
PROJECT OVERVIEW ....……7
Project Description ....………….7
Goals and Objectives 8
Project Milestones ....……8
Project Constraints ....………10
COST BENEFIT ANALYSIS...10
PROJECT PLAN SET-UP..…..12
REFERENCES…………………………………………………………………………………..14

EXECUTIVE SUMMARY

This business case outlines how the machine Learning Project will address current business concerns, the benefits of the project, and recommendations and justification of the project. The business case also discusses detailed project goals, performance measures, assumptions, constraints, and alternative options.

Issue

An entire vehicle manufacturing involves lot of time in production. Any fault in any one of the parts of the vehicle can be very expensive. In the manufacturing phase of vehicles, the process of identifying the root cause of an issue is lengthy and painstaking. Massive amounts of testing data, sensor measurements, and manufacturer parameters are used in Root cause analysis. Performing with traditional methods can be time consuming and hard. Machine Learning techniques will widely accelerate root cause analysis.

Anticipated Outcomes

Predictive analytics in machine Learning can be used to evaluate whether a flawed part can be reworked or needs to be scrapped**.** The manufacturers of the parts can capture images of each component of the vehicle as it comes off the assembly line, and then automatically run the images through a machine learning model to identify any flaws. Highly-accurate anomaly detection algorithms can detect issues down to a fraction of a millimetre.

The same approach can be used for all component manufacturing as well as throughout the vehicle assembly line. Anomaly detection algorithms can analyse vast amounts of system and driver data efficiently. And they can perform this analysis using additional data types and in far greater quantities than traditional methods can handle.

Machine learning can provide far more precise and importantly evolving maintenance recommendations to help drivers protect their vehicle investment as well as their safety.

Recommendation

Accurate datasets containing hundreds or even thousands of parts images, each one tagged with information are required in order to implement an image recognition and analytics model, such as pass, fail, issue A/B/C, etc. The data scientist working on the construction of the model must also have prominent expertise in allowing tolerances and the potential performance and mitigation of impact of various flaws. Image recognition and analytics models can play many roles in the automotive value chain like identifying and analyzing the patterns to help in development of new and better performing tires, a proper control over quality of paint and other finishes and avoiding hazards in Automated Driving Systems.

Justification

With many issues rising in Machine Learning text recognition and Natural Language Processing allows the inclusion of service provider notes in the process of analysis. And all of the processes can help in identifying the root causes in a much faster way than the usual traditional way. Also they identify the problems which cannot be recognized by traditional methods.

Machine Learning can provide accurate and more precise solutions in helping the drivers in protecting their vehicles investment and also the safety of the drivers. The predictive analytics model can continue to learn from its own performance data collected from the manufacturing processes and vehicles unlike the traditional methods which are only updated a few times in a year.

Market Analysis

The global machine learning market is expected to grow from USD 1.41 Billion in 2017 to USD 8.81 Billion by 2022, at a Compound Annual Growth Rate (CAGR) of 44.1% [6]. The main factors for the market are proliferation in data collection, generation and machine learning technological advancement.
In the services segment, the managed service segment is estimated to grow at a higher CAGR, whereas professional service segment is assumed to be a bigger contributor during the estimate time frame. The managed service is supposed to be becoming faster, as it assists associations with expanding productivity and save costs for managing on-demand machine learning services. The development of the professional services segment is mostly governed by the complexity of operations and increasing deployment of ML solutions[5].

The ML in Auto mobility sector report examines top consumers and producers, and focusing on product capacity, consumption, value, market share and growth opportunities in these key areas.

- North America : United States, Canada and Mexico
- Europe : Germany, France, UK, Russia and Italy
- Asia-Pacific : China, Japan, Korea, India and Southeast Asia
- South America : Brazil, Argentina, Colombia
- Middle East and Africa : Saudi Arabia, UAE, Egypt, Nigeria and South Africa

Market Requirements

- **People:** Every ML solution is designed, built, implemented, and optimized by a team of highly trained experts: Machine Learning scientists, applied scientists, data scientists, data engineers, software engineers, development managers, and technical program managers. These skill sets can cost the business project development.

- **Time:** Conducting a machine learning project requires a significant amount of time, from a few weeks to several years. The team of machine learning will have to collect and clean data; design, build, test, and optimize the ML solution. All of these tasks require a serious time commitment and deadline[6].

- **Budget:** Machine learning is a kind of premium service. The company should be financially solvent to hire and "upkeep" the machine learning team for at least several months.

PROBLEM DEFINITION

Problem Statement

The Auto mobility sector has suffered a great loss at the manufacture of defective products. The cost of reproducing or replacing a single component or part of the vehicle or an entire component is very time consuming and highly expensive. The traditional methods of analysing the vehicles before manufacturing is a tedious process. In the manufacturing phase, discovering the main cause of one or many issues is not always possible. Performing these techniques with traditional methods is incredibly hard.

Organizational Impact

This project will impact the auto mobility in several ways. The following will provide the explanation of how the organization, tools, processes and roles and responsibilities will be affected as a result of Machine Learning Implementation.

Tools: The analysis tools used for the predicting the defective components needs to be implemented.

Large datasets which contains parts information is required.

Roles and Responsibilities: In addition to automotive engineering knowledge, the employees should have knowledge about using the ML analysis tools and basic idea about the Ml algorithms.

The new platform will be managed by the IT group and we do not anticipate any changes to IT staffing requirements.

Hardware/Software: In addition to the software and licensing for the project, the company will be required to purchase additional servers to accommodate the processing of large data sets.

Technology Migration

In order to effectively migrate existing analysis on our new platform of Machine learning in auto mobility large data sets of component images needs to obtained and stored. Existing analysis tools to be replaced with the machine learning models.

Phase I: All necessary datasets will be collected. Hardware/Software will be purchased and the system will be created and tested by the IT development group.

Phase II: IT group will stand up a temporary legacy platform in the technology lab to be used for day to day operations for payroll and administration activities. This will be used as a backup system and also to archive all data from the company mainframe.

Phase III: All employees will receive training on the new analytics model.

Phase IV: The analysis tool to be deployed to the system

PROJECT OVERVIEW

The overview consists of a project description, goals and objectives for this project of implementing machine learning in auto-mobility sector, project performance criteria, project assumptions, constraints, and major milestones. As the project is approved and moves forward, each of these components will be expanded to include a greater level of detail in working toward the project plan.

Project Description

The Project will review and analyse several potential algorithm for the implementation. This will be done by collecting large image sets of components and parts of the vehicle manufacturing process. Once collected the project will implement a machine learning algorithm to run the images through it and create an analysis model.

This project will result in greater efficiency in analysing the defective parts before undergoing the whole manufacturing process. Additionally, the engineers in auto mobility will have an overlook on the data modelling and analysing. Once the product is acquired, all implementation and data population will be conducted with internal resources.

Goals and Objectives

The Project directly supports several of the corporate goals and objectives established by the company. The following table lists the business goals and objectives that this Project supports and how it supports them:

Business Goal / Objective	**Description**
Timely and accurate reporting	Analysis tool will allow us to reduce the cost of the traditional methods and saves time in redoing the manufacturing
Improve staff efficiency	Fewer engineers will be required to overlook the process of analysing

Reduce employee turnover	Automating the analysis of the components will result in cutting down employee costs

Project Milestones

Phase 1 – Discovery Stage (7 to 14 days or 1 to 2 weeks)

The machine learning project or product application roadmap begins with the definition of a problem. Firstly, it looks into the issues and operational inefficiencies which should be addressed.

The goal here is to identify the requirements. The stage requires our developers or engineers to meet with the business people, customers on the client side to understand their vision, specifications and requirements in terms of what issues they are looking to solve.

Secondly, the development team should identify which kind of data or datasets they have and if they would need to fetch it from outside service or source (i.e. Kaggle).

Finally, developers have to gauge if they are able to supervise the ML algorithms if it returns the correct response output every time a prediction is made.

Deliverable: A Problem Statement which would define if a product is trivial or would be complex.

Phase 2 – Exploration Stage (4 to 6 weeks)

The main objective of this phase is to build upon a Proof of Concept (PoC) which can then be installed as API. Once a baseline model is trained, the machine learning professionals team estimate the performance of the production-ready solution.

This phase gives us the clarity on what performance should be expected with the metrics planned at the discovery stage.

Deliverable – A Proof of Concept

Phase 3 – Development Stage (3+ months)

This is the stage where the team works iteratively (Agile Framework and Scrum Lifecycle) till they reach a product solution. Because there are far less uncertainties by the time the project reaches this phase, the estimation gets very precise.

But in case the result is not improved, the developers team would have to apply a different algorithm, model or rework on the data or even change the method, if needed.

In this phase, our developers work in sprints (for example, sprint 1, and sprint 2 and so on) and decide what is to be done after every individual iteration in the daily scrum meetings of 15 minutes. The outcomes of every sprint can be predicted effectively.

While the sprint outcome can be predicted effectively, planning for sprints in advance can be overcome the mistakes of methods or algorithms in case of ML.

Deliverable – A production ready ML solution (product or software solution)

Phase 4 – Improvement Stage (continuous till the end of the product)

Once deployed, decision makers are almost always in a hurry to end the project to save budgets or costs which is required for project production or development. While the formula works in almost 80% of the projects, the same does not apply in ML applications.

What happens is that the data or information changes throughout the ML project timeline. This is the main reason why an ML model has to be monitored, reviewed and upgraded continuously – to save it from degradation and provide a safe ML enabling system.

The ML centred products require time and cost for achieving satisfying results. There is also a chances of the program or machine learning script might get lost when any algorithms is used on a different datasets.

Project Constraints

The following constraints apply to the incorporating machine learning in auto mobility. As project planning begins and more constraints are identified, they will be added accordingly.

1. There are limited data for the images of the automobile parts and the resources for machine learning are very limited

2. As implementation will be done internally by the Ml developers and not by the product developers or vendors, there will be limited support from the hardware/software providers.

Cost Analysis :

Action	Action Type	Description	First Year Costs
Data Collection, Training Data Quality and Quantity, Validating, Cleaning and Labelling the Data Samples	Cost	Depending on the nature of the data, the amount of datasets (small, medium and large size), and the complexity of the annotations.	$10,500 - $85,000
Analysis and Research Phase	Cost	Teams have, on average 5 members. Out of these 5 maybe 3 are outsourced (either services or freelancers).So that's 2 employees (2 x $5,000), 3 freelancers (3 x $3,000) resulting in $19,000 per month. Thus, cost is around $ 2,28,000 (12 x $19,000).	$2,28,000
Virtual CPUs and GPUs are running on nodes	Cost	Nearly 4 virtual CPUs running on 1 to 3 nodes for $100-$300 per month. So it costs around $ 3600(12 x $300).	$1200-$3600
latency free machine learning inference	Cost	Spectrum latency cost about $2,500 per month and for year, it cost $ 30,000(12 x $ 2,500).It varies depending on the	$10,000 to $30,000

		complexity of the algorithm being deployed.	
Virtual Machine	Cost	It costs approximately $ 333 per month. Therefore, estimated cost of $4000(12 x $333) for a year.	$4,000
modify the rest of the system to use the new API	Cost	$1,500 for preparing machine learning models to be served and writing the scaffolding for the API for the development. Cost= $18,000(12 x $1,500)	$18,000
solid data pipeline	Cost	$30,000(12 x $ 2,500) per year that includes implementation and documentation.	$30,000
Mmaintenance Cost	Cost	ML model has to be monitored, reviewed and up graded Continuously to save it from degradation and provide a safe ML enabling system.	$37,000
Back End Developer	Salary	To handle database, APIs and other services.	$120,000
Machine Learning Developer	Salary	A ML **developer** is an expert on using data to training models. The models are then used to automate processes like image classification, speech recognition, and market forecasting.	$1,34,000
Total Cost For Annual Year			$ 6,89,600

ML Project Plan Setup :

- Define the task and requirements
- Identify the project feasibility
- Discuss the general model tradeoffs
- Create a project design and plan

Collection and Labeling of Data

- Create the labeling documentation
- Build the data ingestion pipeline
- Validation of data quality

Model Exploration

- Establish the baseline for model performance
- Create a simple model with initial data pipeline
- Try parallel ideas during the early stages
- Find the machine learning model for the problem domain, if any, and reproduce results.

Refinement of Model

- Do model-centric optimizations
- Debug models as complexity gets added
- Conduct error analysis for uncovering failure modes.

Test and Evaluate

- Evaluate the model on test distribution
- Revisit the model evaluation metric, ensuring it drives desirable user behavior.
- Write tests for – model inference function, input data pipeline, explicit scenarios expected in the production.

Deployment of Model

- Expose the model through Restful API
- Deploy the new model to a subset of users to ensure that everything is smooth before the final rollout.
- Have the ability to roll back the models to its previous version
- Monitor the live data.

- Upload the code on GitHub for version control and JIRA for product management.

Model Maintenance

- Retrain the model for preventing model staleness
- Educate the team if there is a transfer in the model ownership

References

[1] Incze, R. (2019, September 14). *The Cost of Machine Learning Projects*. Medium. https://medium.com/cognifeed/the-cost-of-machine-learning-projects-7ca3aea03a5c#:~:text=data%20science%20team.-,Total,is%20a%20very%20optimistic%20estimation.

[2] *Job Bank Canada - Jobs in Canada: Job Bank Canada .com, Find Jobs in Canada Now!* Jobs Bank Canada. (n.d.). https://jobbankcanada.com/.

[3] Indeed. (n.d.). https://www.indeed.com/.

[4] *Glassdoor Job Search | Find the job that fits your life*. (n.d.). https://www.glassdoor.com/Job/index.htm.

[5] Srivastava, S. (2021, June 18). *How to Estimate the Time, Cost, & Deliverables of an ML App Project*. Appinventiv. https://appinventiv.com/blog/machine-learning-app-project-estimate/.

[6] *Machine Learning Market*. Market Research Firm. (n.d.). https://www.marketsandmarkets.com/Market-Reports/machine-learning-market-263397704.html.

[7] *4 Machine Learning Use Cases in the Automotive Sector*. Anaconda. (n.d.). https://www.anaconda.com/blog/4-machine-learning-use-cases-automotive.