

ISYS3420

Machine Learning for Decision Makers

Assignment 3 - Business Pitch

(Transcript)

Rokshana Pervin – s3923324

Video Presentation Link:

https://drive.google.com/file/d/12IpxhoPZd4N4Ia9o5YGenXxAnjnydp7I/view?usp=drive_link

Hi, I am Rokshana Pervin. Today, I am going to represents a Business Pitch in regards of hedge fund business problem and its objective as well as how to solve this problem using machine learning solution.

In the previous hedge fund project, an accurate prediction was not found from the model which results in the business problem and didn't meet the requirements of the hedge fund. As the model prediction were not accurate and share growth were vulnerable, many company refused to invest their profit. In this case to increase the profit from the investment by 10 percent across the year, it is considered to reduce the number of companies whose values falls by a certain threshold value which is 20 percent.

To solve this problem, machine learning approach will be used by following Data science lifecycle where business understanding, data understanding, model evaluation is included.

In this case, the different components of the solution would be used such as reformulating the business question where the number of companies whose values falls by 20 percent will be reduced to get the strategic goal. While revisiting the data exploration and analysis, the threshold value will be reviewed which could help to solve the business problem and met the requirement.

The organization's benefit is that the profit will increase by 10 percent across the next year.

If we get this benefit, hedge fund will earn their revenue across the next year.

As organisation will grow with their revenue, revenue need to calculate using ROI calculation or Return on Investment calculation that can estimates the expected monetary value and will reduce the risk of investment. Besides this, when organisation will earn more revenue, more investors will be attracted. They will grow their trust on their investment. In this case, we can use one simple mathematical equation:

$$\text{returns} = \text{value} - (1 - \text{accuracy}) * \text{cost of a mistake}$$

Here, the variables are,

- **returns:** Generated net value or profit per prediction.
- **value:** The new value generated by every prediction (e.g. assigning a document to the right category now takes 0.01 seconds instead of 5 minutes, so the value is 5 minutes saved)
- **accuracy:** The accuracy of predictions made by the algorithm.
- **cost of a mistake:** The additional costs incurred by a wrong prediction (e.g., it takes 20 minutes for someone to correct the mistake in the system)

As organisation will grow with their revenue, revenue need to calculate using ROI calculation or Return on Investment calculation that can estimates the expected monetary value and will reduce the risk of investment.

In this case, we can use one simple mathematical equation:

Here, above these contexts are related to the machine learning approach. We can use these contexts based on the previous machine learning project. When we will arrange all of the values of above features by operating the machine learning algorithm, we can calculate how much profit we will get from the machine learning

deployment, or we will understand where we should need to pay attention in terms of cost of a mistake.

In choosing machine learning solution, trade-offs need to consider where cost of technical expertise, cost of machine learning maintenance are included. We need to consider several machine learning risks such as malicious input that can cause from false predictions, need to consider the fraction of the part of the data that can be attacked by the attacker in a concerted system and monitoring the online system. Apart from this, another challenging risk in the machine learning system can occur from data breaching or breaching the data confidentiality. Hence, we need to count all of the aspects of the machine learning process in terms of addressing the problem or opportunity.

After observing the project with the cost, benefit, and risk, we need to care about another challenge known as model drift or the `concept drift` that occurs from the changeable data streams over the time in the real world which results the less accurate predictions for the past trained model. As a result, it indicates to continue the machine learning development or testing to extend the opportunities of the algorithm. Here, I am proposing to implement my machine learning result where metric selection and changeable threshold value can help to fulfill the business requirements as well as to reduce the risk of investment. After implementing the machine learning results, the purpose will be:

- Predicting True Positive (TP) more
- Reducing False Positive (FP) more.

Because, when the previous threshold of the business case was 0.2, it couldn't predict the profit or revenue for the organisation adequately which results the risk of investment. After changing the threshold value according to the business requirements to reduce the risk of investment, the following results have been obtained where recall and F1-Score are highlighting the significant output for the GradientBoosting Classifier compared to other metrics. Here, recall is the calculated proportionate term which helps to identify the actual positive value that were correct. In this business case, as we need to reduce the riskier companies in terms of their falling invested values, therefore, F1-Score will help to take the decision to organise the risk-free companies as F1-score or F-measure is in combination of precision (predicted values) and recall (actual values) terms where both deals with true positive compared to other metric.

Models Name	Results/Output	Thresholds	Precision	Recall	Accuracy	F1-Score

GradientBoosting Classifier	Previous	+0.2	0.66	0.66	0.74	0.66
	Present	-0.2	0.82	0.98	0.81	0.89

$$\textbf{Precision} = \frac{TP}{TP + FP}$$

$$\textbf{Recall} = \frac{TP}{TP + FN}$$

$$\textbf{F - measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

In this way, the implementation of the machine learning solution will be aligned to the strategic direction of the hedge fund organisation where hedge fund will be benefited by reducing the riskier companies whose values falls by a certain threshold and will earn profit from the investment by 10% across the next year.

Apart from this, to manage, govern, and maintain this solution in production or organisation, the machine learning principles or MLOPs principles will be followed.

To achieve the ultimate vision alongside the machine learning work and to improve the machine learning solution in the future, I will implement the MLOPs principles by following the best practices and tools where validating the MLOPs process or machine learning operations processes are included. The MLOPs processes are represented with business understanding, defining compatible machine learning algorithm as well as testing, continuous delivery, and monitoring.

To make understand the knowledge of the machine learning algorithm and to instill confidence in the algorithm in decision-makers, the decision-makers will be continuously guided by me by following the MLOPs principles so that decision-makers can take their right decisions for their organisations.

In this case, the right time to expose the solution to the decision-maker would be the early stage in the machine learning process so that decision-makers or stakeholders can understand the end-to-end process of the machine learning solution.

- Obtained the true positive values using F1-Score which is combined with predicted values(precision) and actual values(recall).

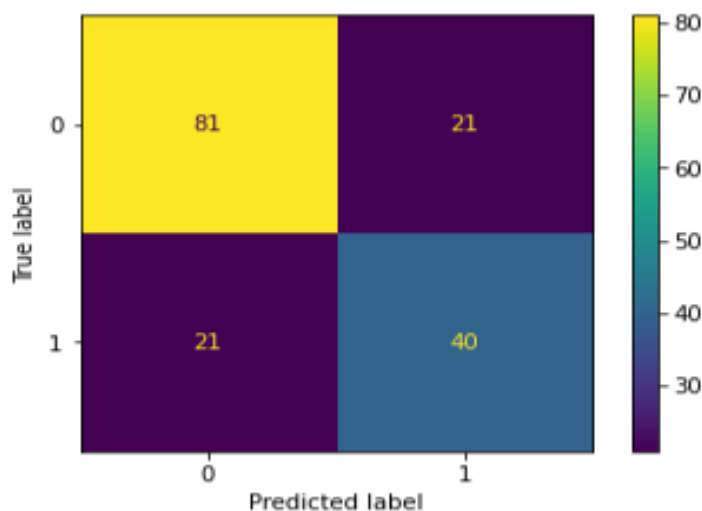
- Reduced the risk of investment by removing the riskier companies whose values falls by a certain threshold value.
- Attracted more investors by showing the profit growth.
- Increased business growth, size, and revenue streams.
- Adapt more easily to changing market conditions as the above machine learning solution (Gradient Boosting classifier) learn gradually from its change.
- Improve business operations.

I will help the decision-maker to realise the benefit by applying data visualisation. Here, according to the threshold +0.2, my goal was to increase the share price above twenty. Therefore, the following binary number 1 referred to grow the share above twenty whether 0 indicated share price growth less than twenty. In this case, I had needed more actual true positive values whereas I got only TP = 40 for +0.2 threshold value which didn't fulfill the business requirements.

On the other hand, I have used the threshold value -0.2 which means it will help to reduce the riskier companies whose values falls by a certain threshold value that is 20%.

As a result, it is proof that if hedge fund manager takes this machine learning result and implement it in the hedge fund, the risk of investment will reduce more and will earn more trust from the investors which will generate the risk-free business and profit from the investment across the next year.

Moreover, if hedge fund uses this result by following the return on investment or ROI calculation, they will be more benefited while they will use `F1-Score` instead of `accuracy` as F1-Score is the important metric for this model here.

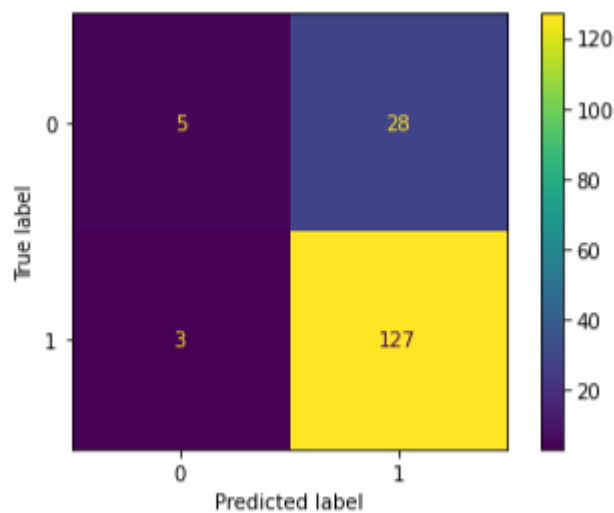


Label

Prediction

Result

0	0	TN = 81
1	0	FP = 21
0	1	FN = 21
1	1	TP = 40



Label	Prediction	Result
0	0	TN = 5
1	0	FP = 3
0	1	FN = 28
1	1	TP = 127

To meet the model's outputs with the business requirement and to reach the business goal, I have used the threshold value -0.2 which means it will help to reduce the riskier companies whose values falls by a certain threshold value that is 20%. Here, the above binary number 1 will be risk free company and 0 will be riskier company. From the above visualisation, it is observed that the actual true positive value is higher that is TP = 127 for this threshold value. As a result, it is proof that if hedge fund manager takes this machine learning result and implement it in the

hedge fund, the risk of investment will reduce more and will earn more trust from the investors which will generate the risk-free business and profit from the investment across the next year.

Moreover, if hedge fund uses this result by following the return on investment or ROI calculation, they will be more benefited while they will use 'F1-Score' instead of 'accuracy' as F1-Score is the important metric for this model here.

Ethical consideration is playing a vital role in use of machine learning algorithms. There are many things can happen in considering the poor business decision that led to make the business down such as a major loss of investment, individuals' reputation as well as monetary loss for the organisation.

Apart from this, during machine learning project, few things can rise such as ethical questions, bias and model drift and model sustainability. If these terms arise during the project or decision making, project will not run with the business. As a result, stakeholders will lose their trust and will refuse to invest in the business market. Eventually, it will increase the customer churn rate.

If something went wrong, the business would go down, even if, it may stop the business. In this case, we need to mitigate some risks and challenges such as bias and the model drift and to increase the model sustainability. We should focus on some stages so that it doesn't happen on the real world. These are:

- Analysing the model using bias-detection techniques.
- Constantly monitoring and retraining the model using new data to reflect with the real world.
- Increasing the robustness which is the ability to continue the model with the best performance.
- Recognising the accountability and responsibility that helps to understand the individuals for a machine learning model's output.
- Considering about risks, security and confidentiality of the machine learning algorithms that can breach by selection and self-selection bias and may lead to a high amount of fine for the organization.

According to the business requirements, we needed to increase the profit by 10 percent across the next year by reducing the riskier companies whose values falls by a certain threshold value. Here, machine learning algorithm precisely met the requirements and acted to solve the problem using the threshold value -0.2 which has been highlighted from the above confusion matrix visualisation.

For implementing the solution, I will offer first short-term plan which can also use as a minimum viable product (MVP) that helps to get the feedback about the product from the customer as quickly as possible. When short-term plan will be successful, long-term and iterative plan will be anticipated to increase the robustness and reliability of the machine learning model. In this case, additional investment can help

to avoid any risk or negative impact around the project. If hedge fund can demonstrate superior investment, it will obviously bring a positive impact to the entire project that can generate the long-term implementation for the solution. Business will be benefited through this additional investment as it may guide the entire machine learning maintenance process such as cost of technical expertise, cost of infrastructure etc as well as it can help to expand the customer retention strategy as customer will attract to invest more after observing the profit, gaining trust from the share market.

After implementing the project, decision maker should care about the project by doing additional research or getting feedback from the customers which can help them to understand the current risks and challenges such as data security and confidentiality. If decision-makers would care about their project from the early stage of the project, they may save huge amount of revenue or profit in every year.

A proof of concept will help to validate and financially viable the machine learning solution which will help to develop the business market rapidly.

I will scale this solution from a small-scale project into an asset across the business by applying both short-term and long-term plan. Because, if the solution develops the decision-makers' confidence and trust while using short-term plan, the plan will transfer into a long-term basis for the business improvement.