



AI & MACHINE LEARNING FOR BUSINESS

Propensity to Buy





TABLE OF CONTENT

01

AI & MACHINE LEARNING

Definition, type, roadmap,
function

02

WHY ML?

Why do we need machine
learning

03

PROPENSITY TO BUY

Prediction of buying
customers in next 5
months

04

NEXT STEP

Future plan





AI & MACHINE LEARNING



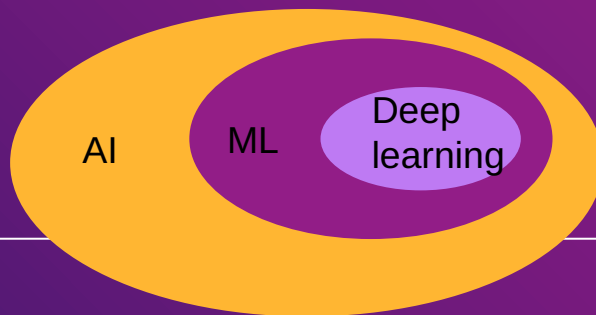
ARTIFICIAL INTELLIGENCE

A technique which enables machine to mimic human behavior.

MACHINE LEARNING

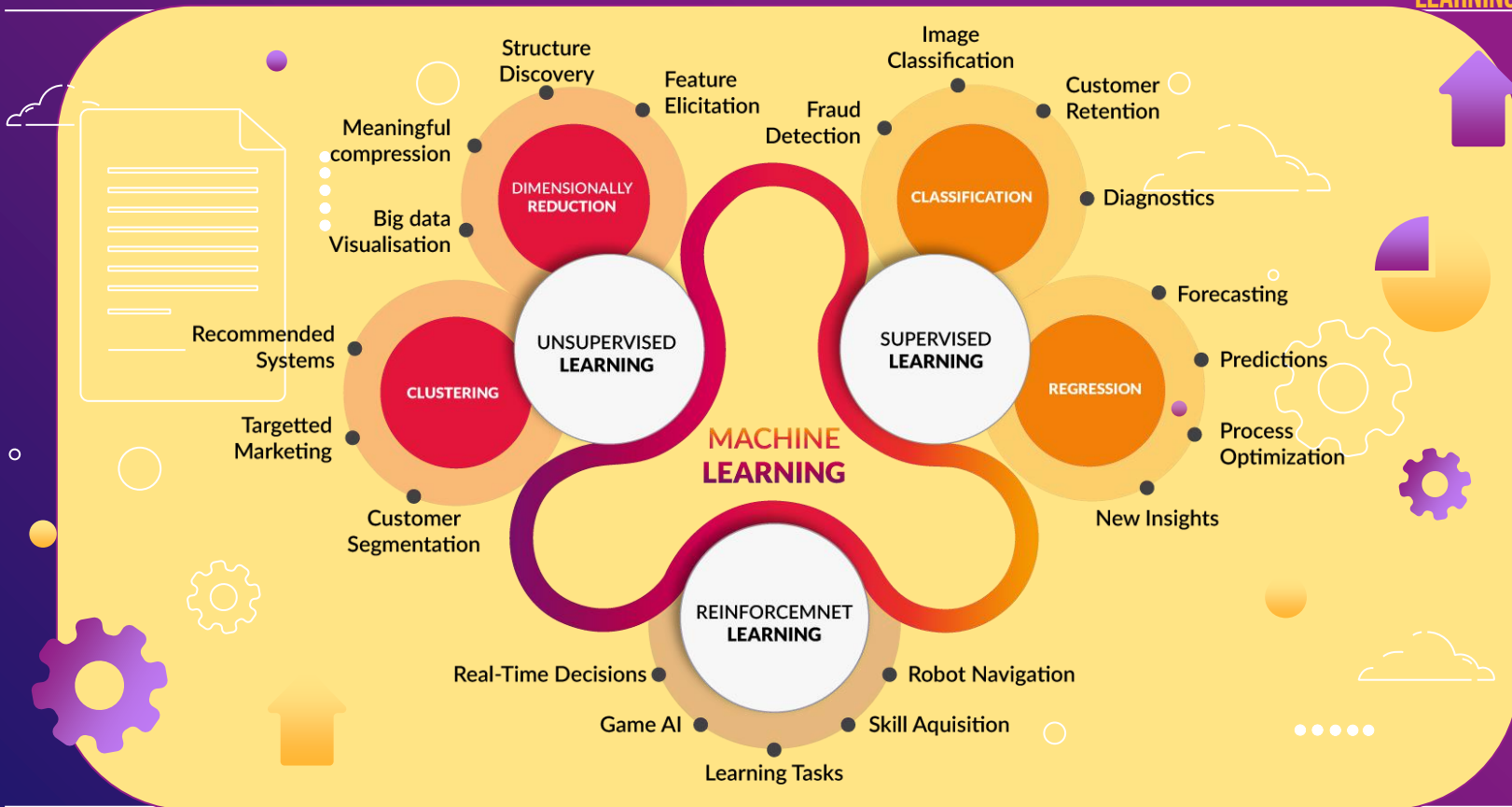
Subset of AI techniques which use statistical methods to enable machines to get better at task without explicit instructions and relying on patterns.

In a relationship:



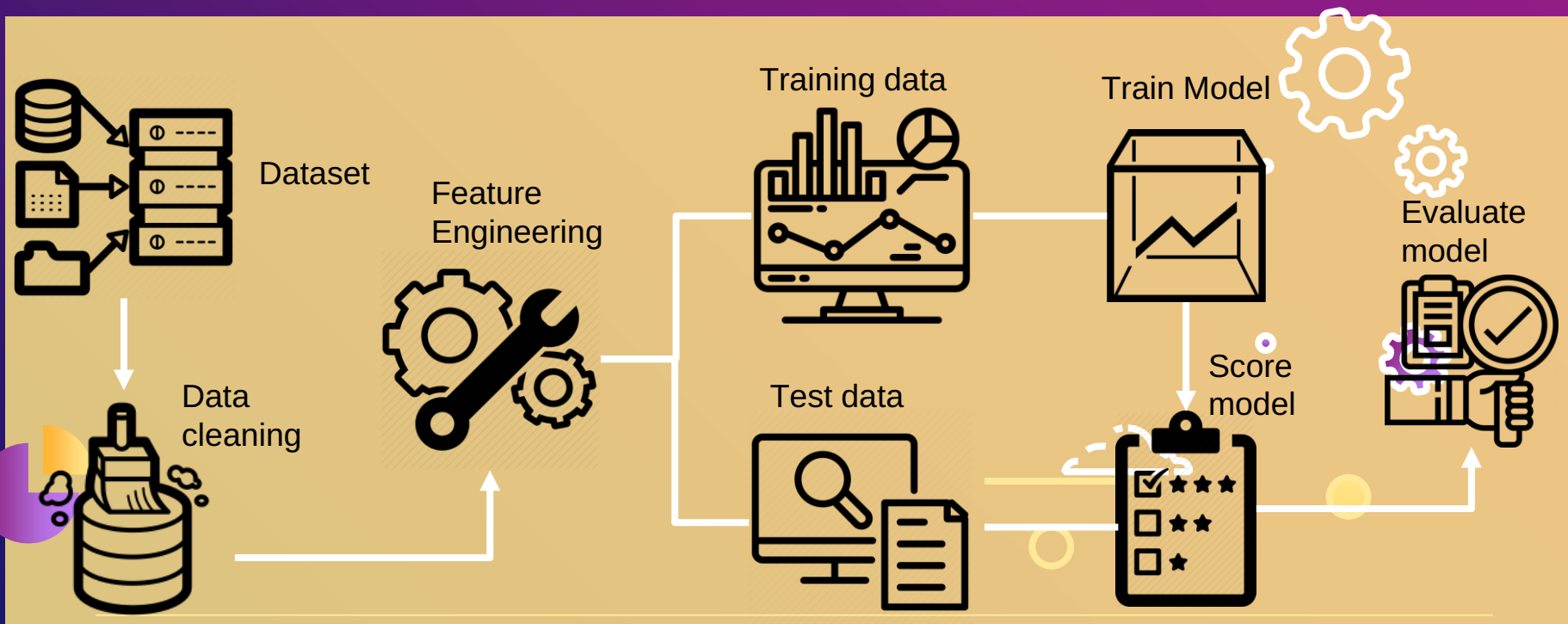


MENU





MACHINE LEARNING ROADMAP





WHY DO WE NEED ML?

- There are some tasks that are too big and too hard for people to do quickly.
- Human can detect pattern, however with the growth of data, it's impossible to detect thousands of pattern from data, not even mentioning the iterations.
- Human are more reluctant to mistakes when doing repeated jobs, while computers are less reluctant to it.





PROPENSITY TO BUY

Here, the propensity to buy model is the predictive model that analyze the purchase behavior of customers by making distributional assumptions

Benefits:

- Predict probability of customer to become churn
- Predict CLV
- Predict purchase volume

Model used: Lifetimes Library (<https://lifetimes.readthedocs.io/en/latest/Quickstart.html>)

Model limitation: only can be used for user with >1 transaction

**Lifetimes use the combination of Beta-Geometric/Negative Binomial (BG/NBD) model combined with the Gamma-Gamma model for estimating the customer lifetime value



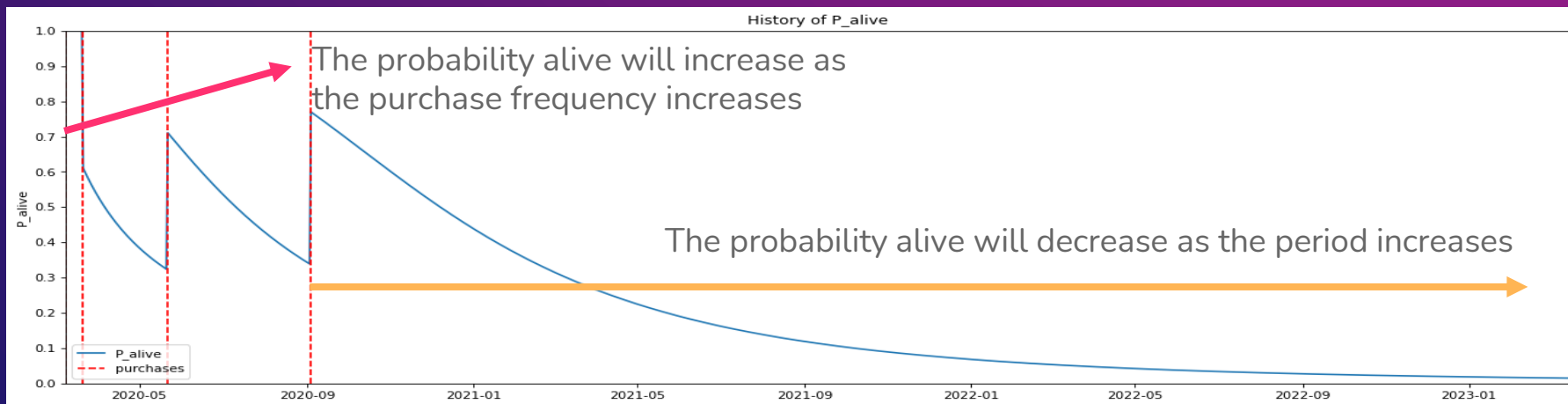
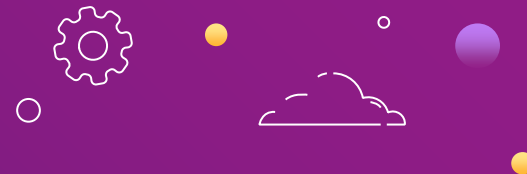


HOW THE MODEL WORKS

The model uses several probability models.

The main factors affecting: the frequency of purchases, the period between every purchase

For example from the user purchase history below, the probability of alive at 1 May 2021 is 0.2



** P_{Alive} is $(1 - \text{probability customer becomes inactive})$, or the probability of the customer is still active in the merchant





PROPENSITY TO BUY MODEL ASSUMPTIONS

Assumption	Explanation
Customer's transaction rate follows Poisson process	The customer's transaction is a random process that is independent and has a mean rate
Heterogeneity in transaction rates among customers follows a Gamma distribution	Each customer has their own probability of buying that follows gamma distribution (a continuous distribution)
The probability of dropping out after transaction follows a geometric distribution	The geometric distribution is used to model the distribution of the number of transactions a customer makes before dropping out to generate a fixed parameter probability p
Heterogeneity in p follows a Beta distribution	Different customer may has different inactive probability (p). The distribution of p across the population of customers can be described by Beta distribution (a continuous probability distribution)
The transaction rate λ and the dropout probability p vary independently across customers.	The expected number of transactions per unit time (λ) and the inactive probability of a customer (p) are assumed to be independent of each other across different customers





DATA DESCRIPTION



Description	Lifetime Transactions from members who joined before 2021
Features	User ID, Transaction recency (date difference between last 2 transactions), Transaction frequency
Train Test Split	The Data was split into two parts, Train (trx date before Dec 2022), Test (trx date on Jan - May 2023)





MODEL EVALUATION



Below is the result for number of purchases in Jan – May 2023 (5 months purchase prediction):

		Predicted Label	
		Positive	Negative
Real Label	Negative	FP = 40,477	TN = 412,517
	Positive	TP = 35,370	FN = 60,309

True Positive Right Number of trx : 17,876
True Positive incorrect number of trx: 17,494

Accuracy (only true positive with right number of trx counted): **78.44%**

Accuracy (All): **81.63%**





NEXT STEP

- Increase model accuracy -> explore other data preprocessing
- Deploy model for production purposes
- Explore more Machine Learning models for business purposes





RESOURCES

- <https://www.getcensus.com/blog/primer-on-the-types-and-features-of-fraud-prediction-models>
- <https://lifetimes.readthedocs.io/en/latest/Quickstart.html>
- https://github.com/CamDavidsonPilon/lifetimes/blob/master/lifetimes/fitters/beta_geo_fitter.py
- <https://towardsdatascience.com/predicting-customer-lifetime-value-with-buy-til-you-die-probabilistic-models-in-python-f5cac78758d9>
- <https://cxl.com/blog/propensity-modeling/#h-1-selecting-the-features-for-your-propensity-model>

Code:

<https://github.com/cindysuyitno/Churn-Prediction>

Thank You!

