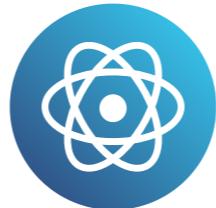


Machine learning mistakes

MACHINE LEARNING FOR BUSINESS



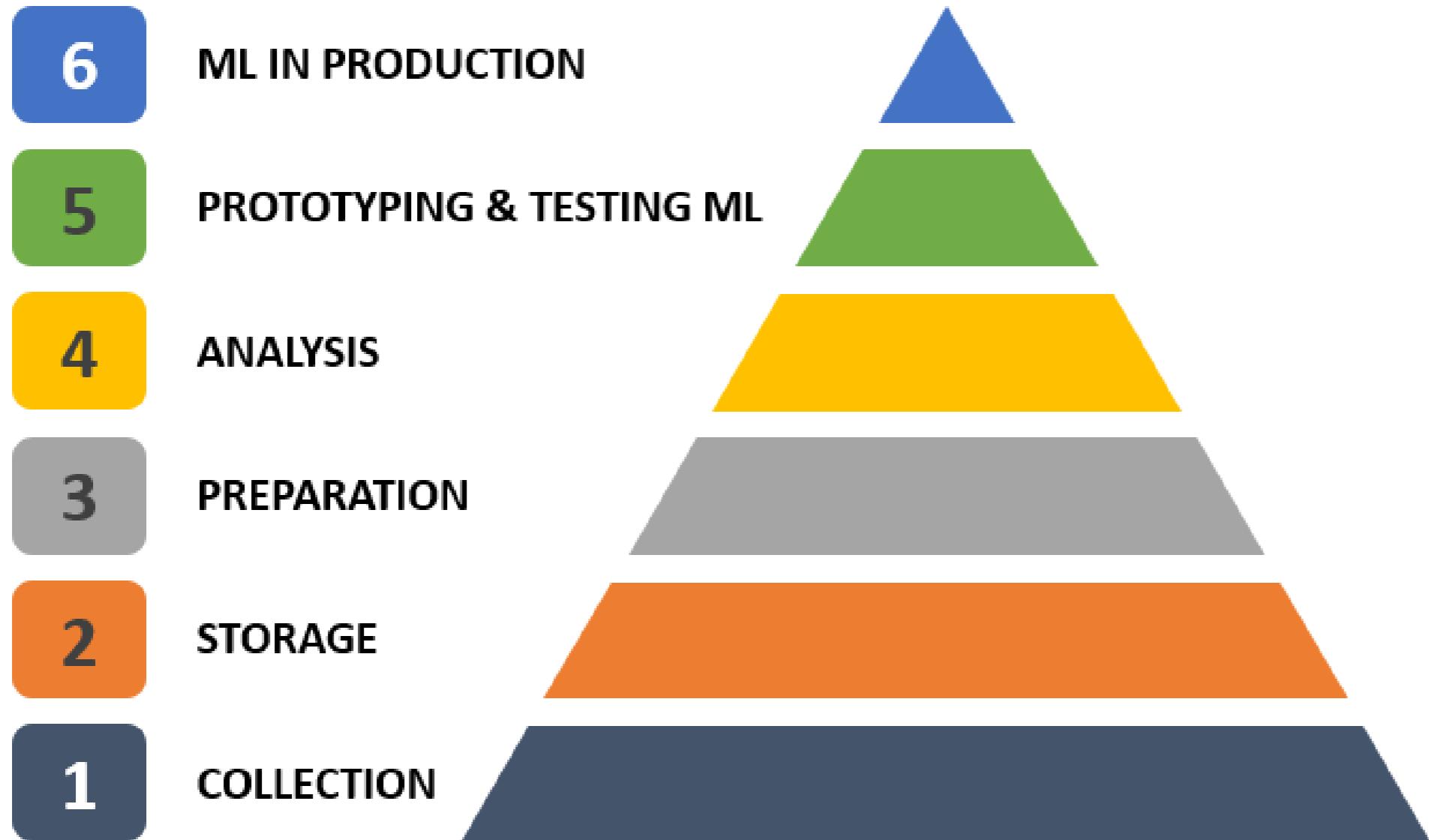
Karolis Urbonas

Head of Machine Learning & Science,
Amazon

Mistakes

- Machine learning first
- Not enough data
- Target variable definition
- Late testing, no impact
- Feature selection

Machine learning first



Not enough data



Target variable definition

- What are we predicting?
- Can we observe it?
 - Contractual churn - customer terminated the premium credit card
 - Non-contractual churn - customer started using another grocery store
- In-depth analysis
- Business field expertise

Feature selection

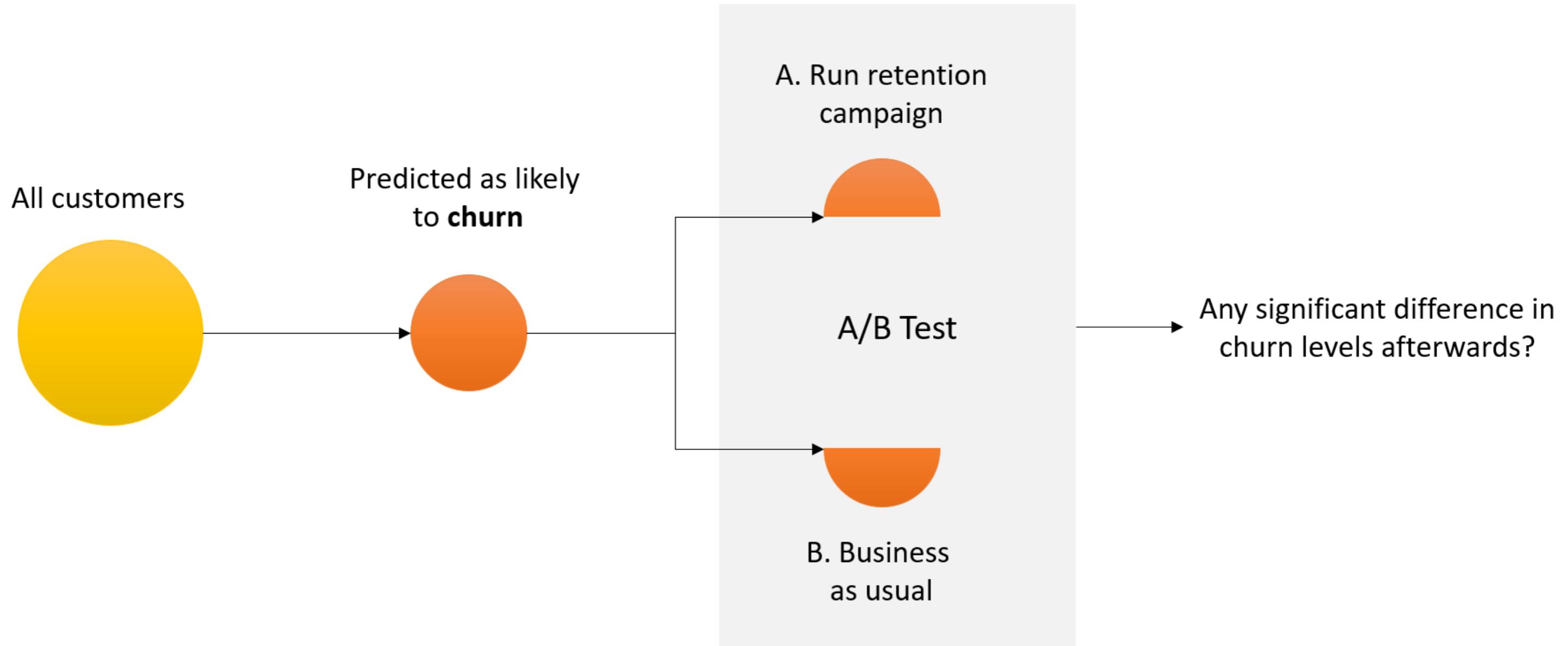
Inference (what affects the target variable?)

- Choose variables that you can control (website latency, price, delivery, customer service etc.)
- Business has to be involved in feature selection

Prediction (can we estimate the target variable value in the future?)

- Start with readily available data
- If model performance is OK, test it
- Introduce new features iteratively

Late testing, no impact

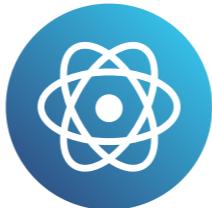


Let's practice!

MACHINE LEARNING FOR BUSINESS

Communication management

MACHINE LEARNING FOR BUSINESS



Karolis Urbonas

Head of Machine Learning & Science,
Amazon

Working groups

Schedule recurring meetings to track progress and define the following:

- Define the business requirements
- Review machine learning model and business products
- Inference vs. prediction
- Baseline model results & outline model updates
- Market testing
- Production

Business requirements

1. What is the business **situation**?
 - *Churn rate has started increasing*
2. What is the business **opportunity** and how big is it?
 - *Reduce churn from X% to Y%*
3. What are the business **actions** we will take?
 - *Run retention campaigns targeting customers at risk*

Machine learning products

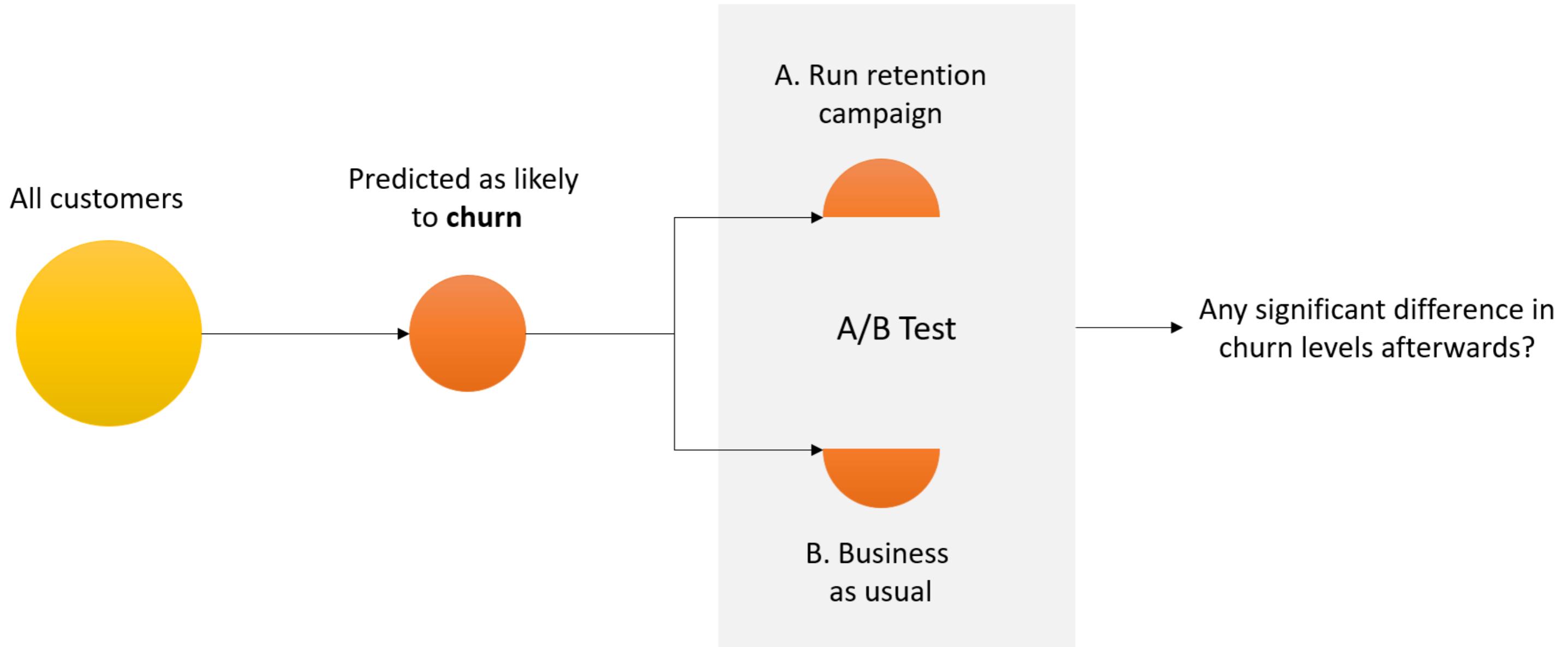
- What ML products does the business needs?
 - **Example 1** - Predict churn. Business wants 1) inference into drivers of the churn updated quarterly, and 2) daily customer classification into: lost customers, customers at risk, no risk
 - **Example 2** - Fraud prediction. Business wants 1) inference into strong indicators of churn, and 2) real-time list of very risky transactions for manual review and medium risk ones for additional data request

Model performance and improvements

Identify what is the tolerance for model mistakes (remember - all models are wrong):

- **Classification**
 - Which class is more expensive to mis-classify?
 - Example - it's likely more expensive to mis-classify fraud as non-fraud than vice versa
- **Regression**
 - What is the error tolerance for prediction?
 - Example - in demand prediction the company will have to buy more inventory than needed if the model error is very high

Market testing



Machine learning in production

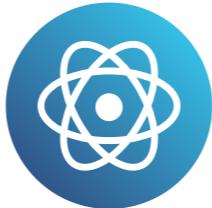
- Are test results delivering consistent positive improvements?
- Is the model stable enough?
- Do we have systems and tools where the model can be integrated to?

Let's practice!

MACHINE LEARNING FOR BUSINESS

Machine learning in production

MACHINE LEARNING FOR BUSINESS



Karolis Urbonas

Head of Machine Learning & Science,
Amazon

Production systems

Production system is live, customer facing and business critical.

- Customer Relationship Management (CRM)
- Fraud detection system
- Online banking platform
- Autonomous cars

CRM

Production system - Customer Relationship Management (CRM)

Example - predicted churn triggers automatic emails



Fraud detection system

Production system - fraud detection system

Example - predicted fraud probability
automatically triggers transaction block and
requests a manual review

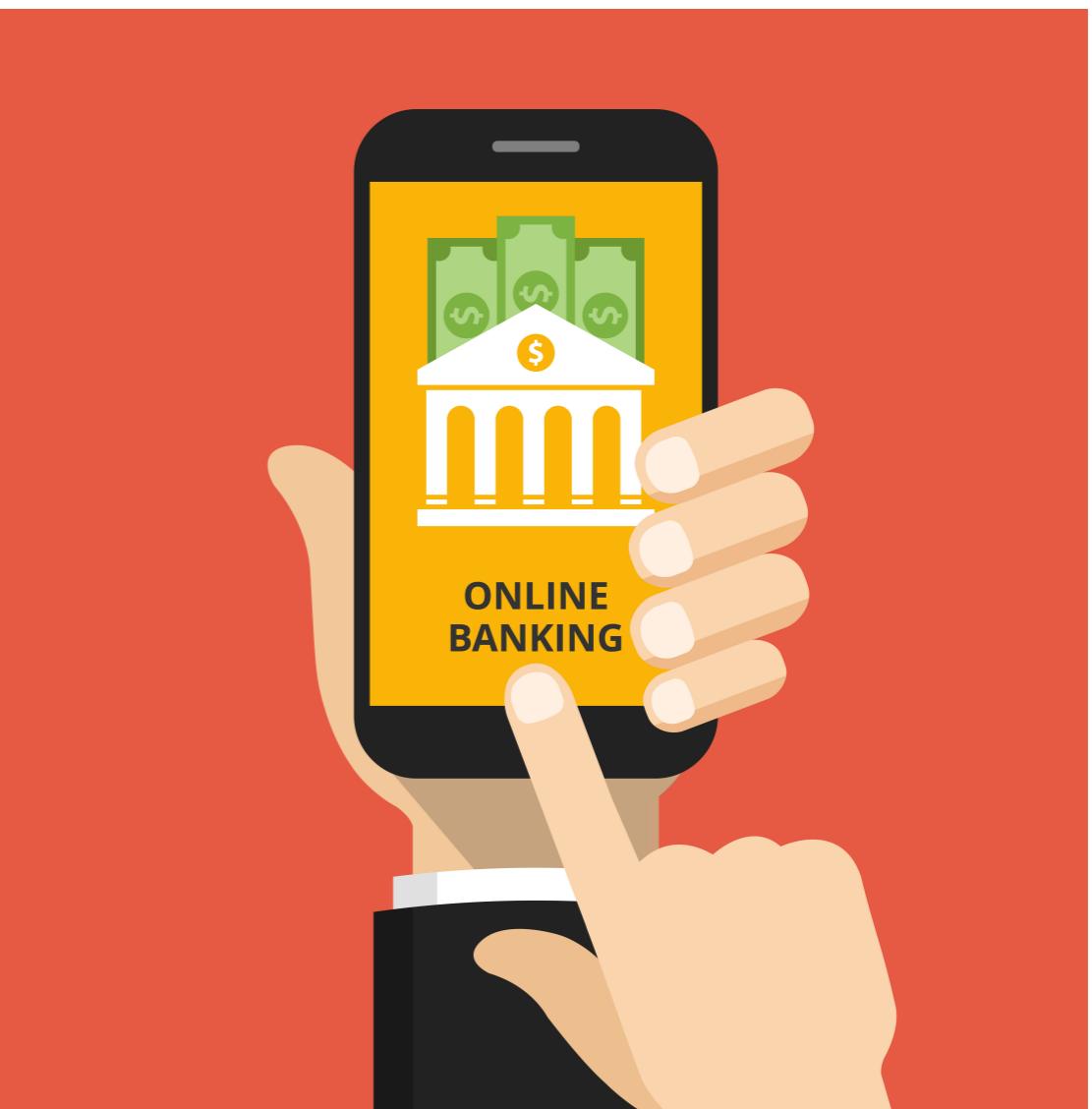


**FRAUD
DETECTION**

Online banking system

Production system - online banking platform

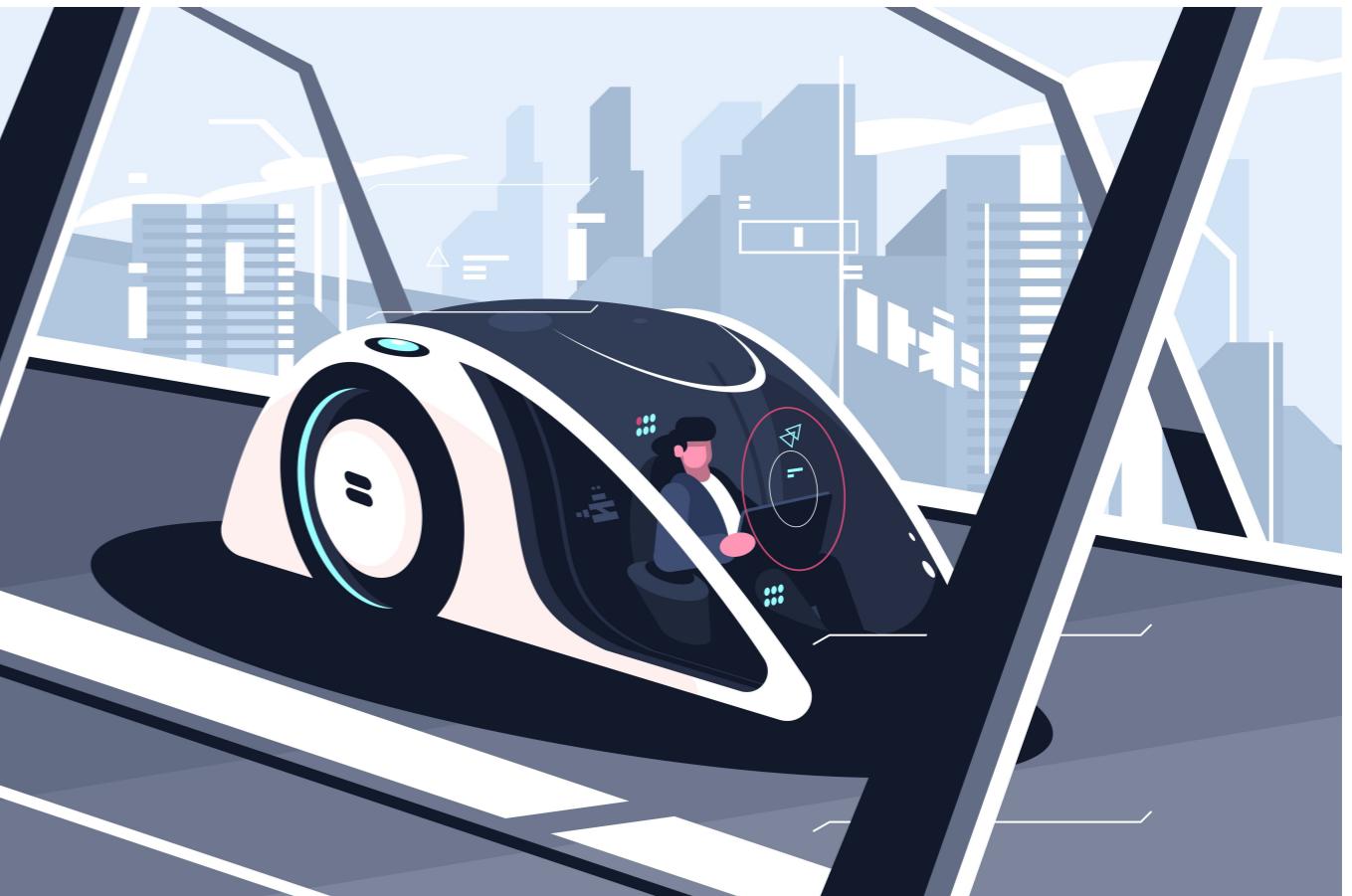
Example - recommended products shown on the customer's online banking profile



Autonomous cars

Production system - autonomous cars

Example - predicted collision kicks off automatic initiation of brakes and collision avoidance steps



Staffing

Prototype ML

- Data Scientists
- ML Engineers

ML in production

- Software engineers
- Data Engineers
- Infrastructure owners

Launch, tracking and feedback

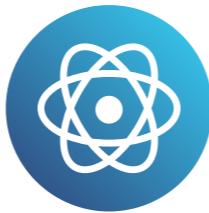
1. Murphy's law
2. Launch to a small subset of customers
3. Track results and if they're consistent
4. Track performance, stability, customer feedback
5. Scale up
6. Repeat 3, 4, 5

Let's practice!

MACHINE LEARNING FOR BUSINESS

Wrap-up

MACHINE LEARNING FOR BUSINESS



Karolis Urbonas

Head of Machine Learning & Science,
Amazon

Thank you!

MACHINE LEARNING FOR BUSINESS