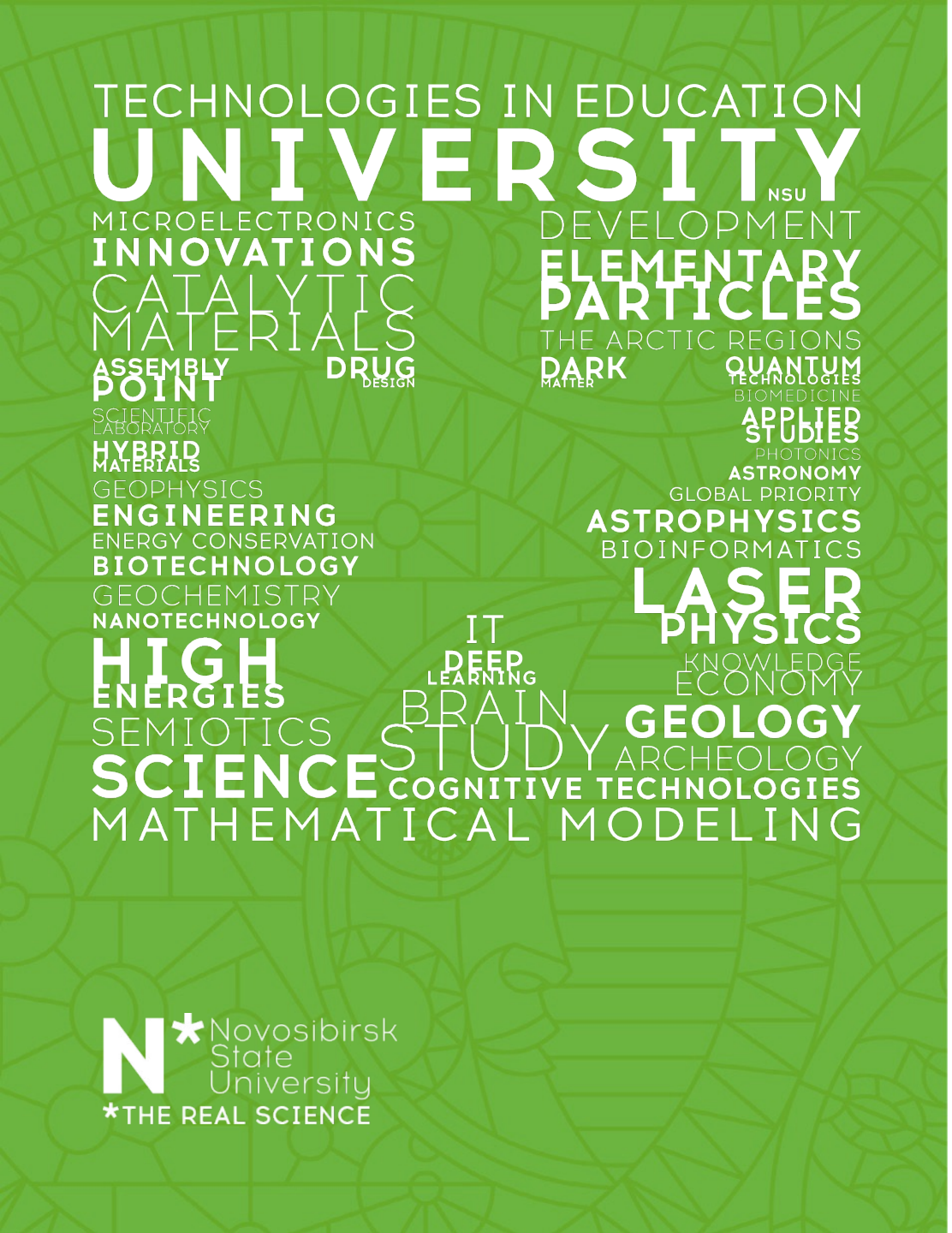


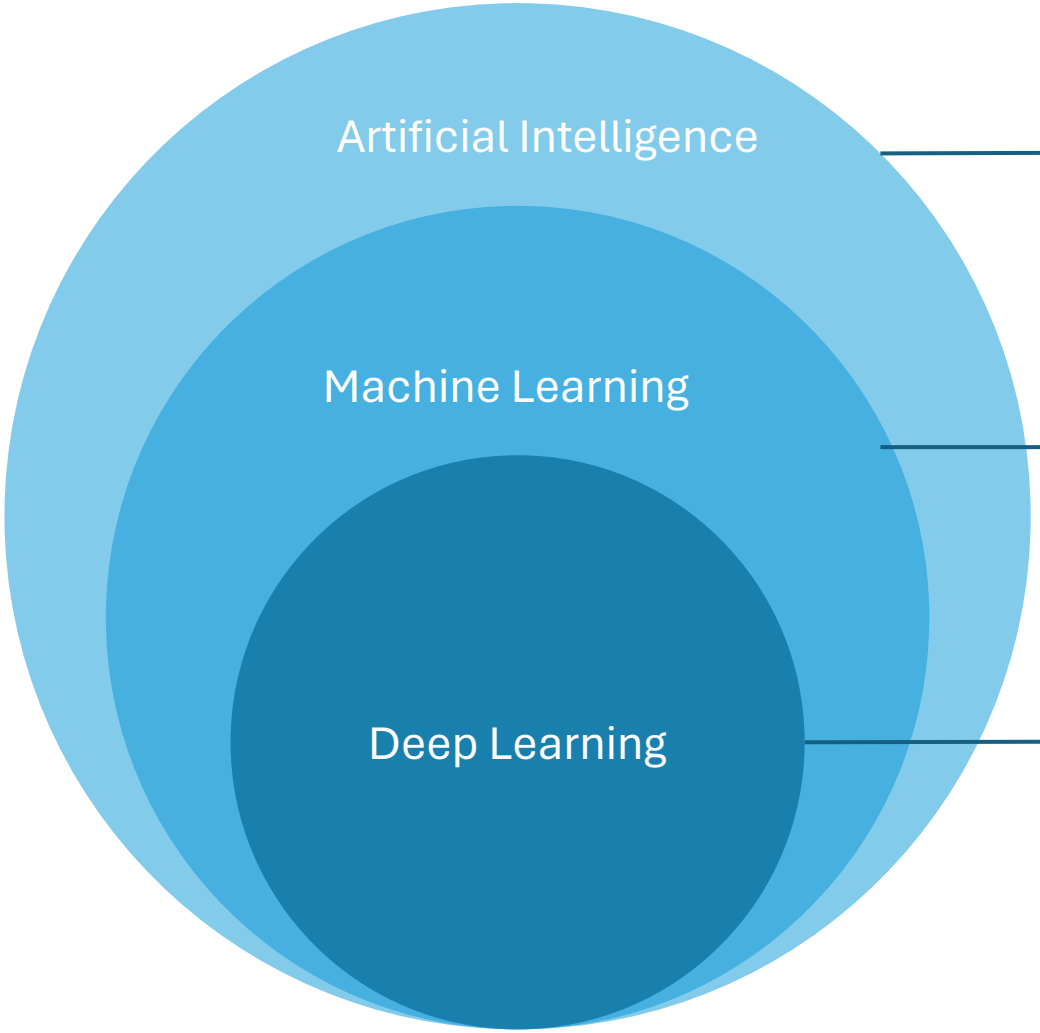
Introduction to Machine Learning

Luu Minh Sao Khue
Department of Mathematics and Mechanics

Course Overview

- Github: https://github.com/luumsk/NSU_ML
- Email: khue.luu@g.nsu.ru
- Telegram: @khueluu
- Notes:
  - Theory + practice (*)
  - Lectures + assignments + personal final project
  - Extra points

What is machine learning?

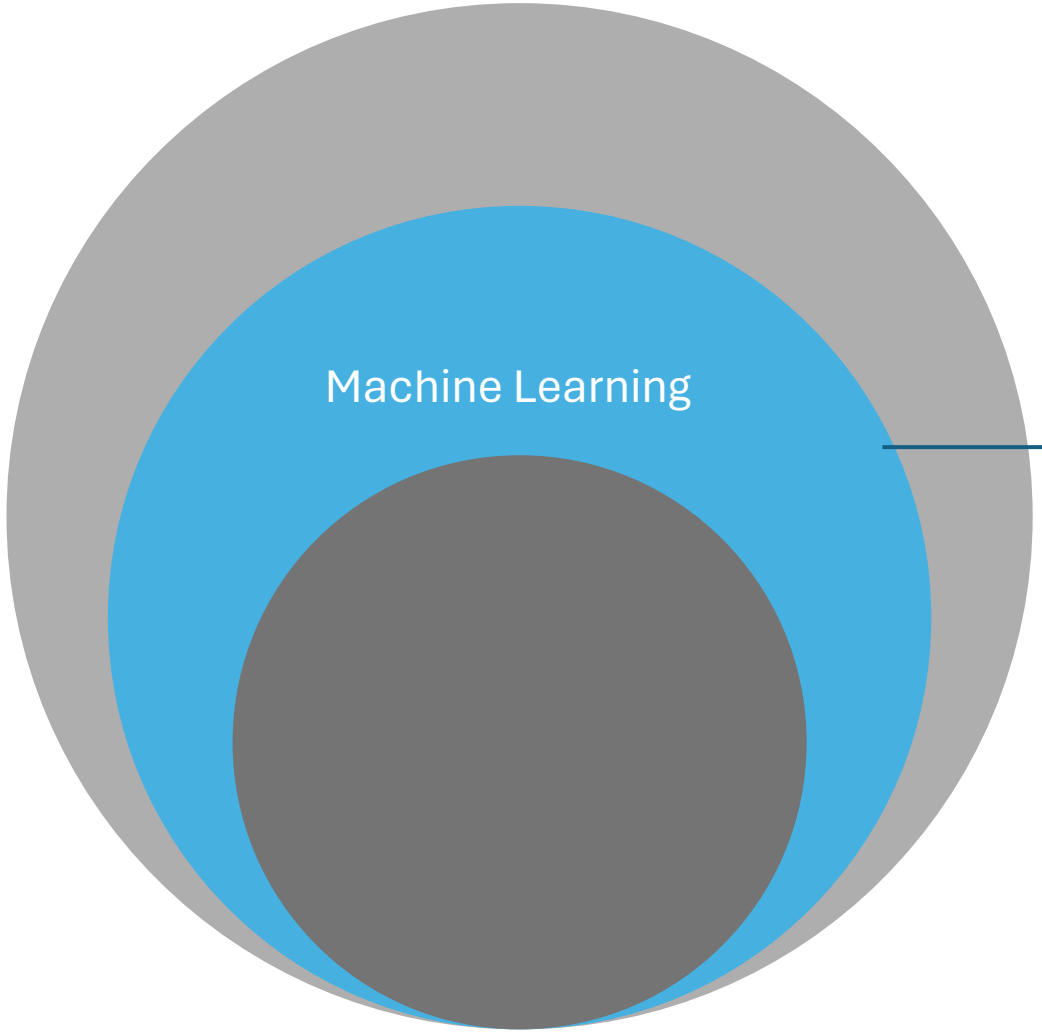


The broad field of creating systems that **mimic human intelligence** to perform tasks autonomously

A subset of AI that enables computers to **learn patterns from data** and improve over **time without explicit programming** for every task

A specialized area of ML that uses **neural networks** with multiple layers to model complex patterns.

What is machine learning?



A subset of AI that enables computers to **learn patterns from data** and improve over time **without explicit programming** for every task

Machine Learning Engineer Salary in Russian Federation

This page is an excerpt of the much more complete compensation information available in ERI's Assessor Series.

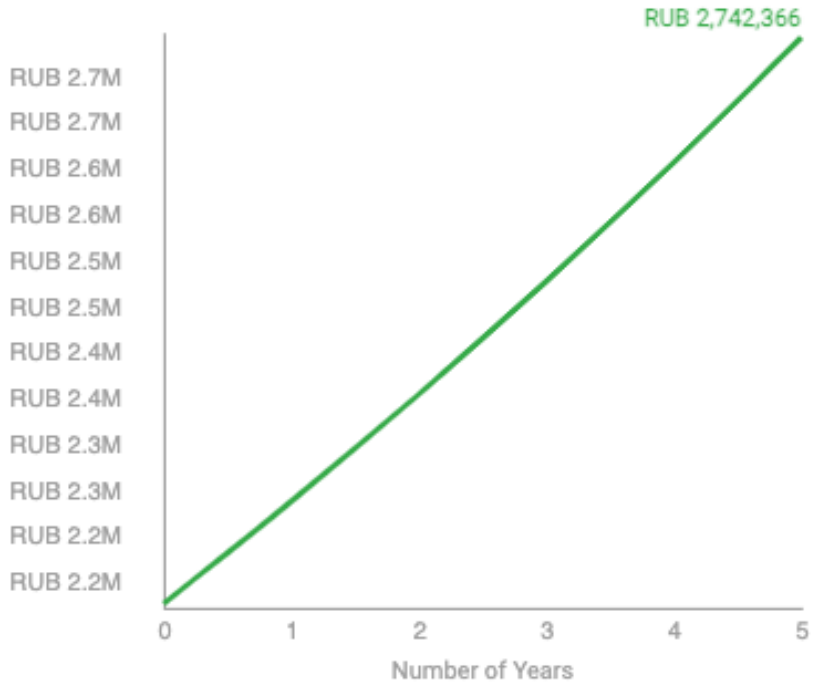
RUB 2,126,200	RUB 1,022/hr	RUB 97,593
Average Salary	Average Hourly	Average Bonus

Estimated salary in 2029:

RUB 2,742,366

5 Year Change:

▲ 29 %

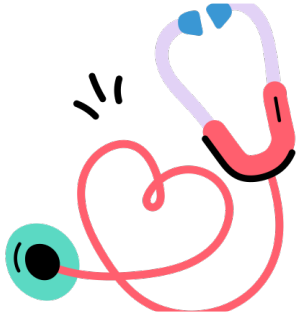


Job titles

- Machine Learning Engineer
- Data Scientist
- Research Scientist (Machine Learning)
- Applied Machine Learning Scientist
- Machine Learning Analyst
- Lead Machine Learning Scientist
- Director of Machine Learning
- Chief AI Officer

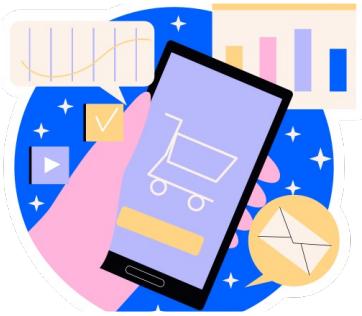
Source: https://www.erieri.com/salary/job/machine-learning-engineer/russian-federation

Applications of machine learning

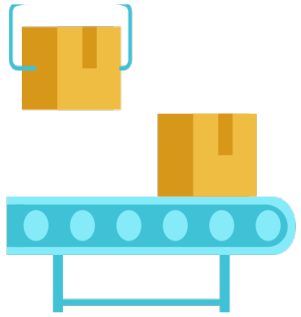


Healthcare

Finance



E-commerce

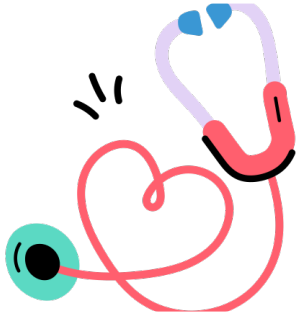


Manufacturing

Transport and Logistic

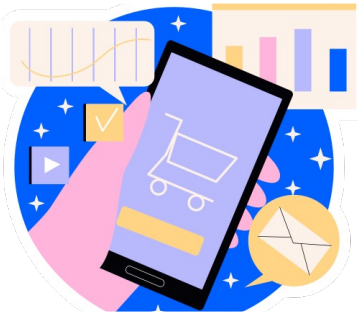
Applications of machine learning

- Diagnostics
- Predictive Analytics
- Personalized Medicine



Healthcare

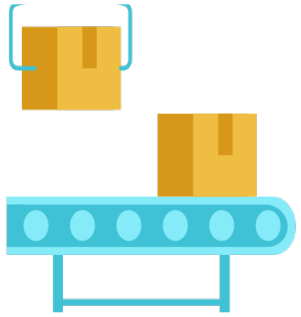
Finance



E-commerce

- Recommendation Engines
- Customer Segmentation
- Inventory Management

- Fraud Detection
- Credit Scoring
- Algorithmic Trading



Manufacturing

- Predictive Maintenance
- Quality Control
- Supply Chain Optimization

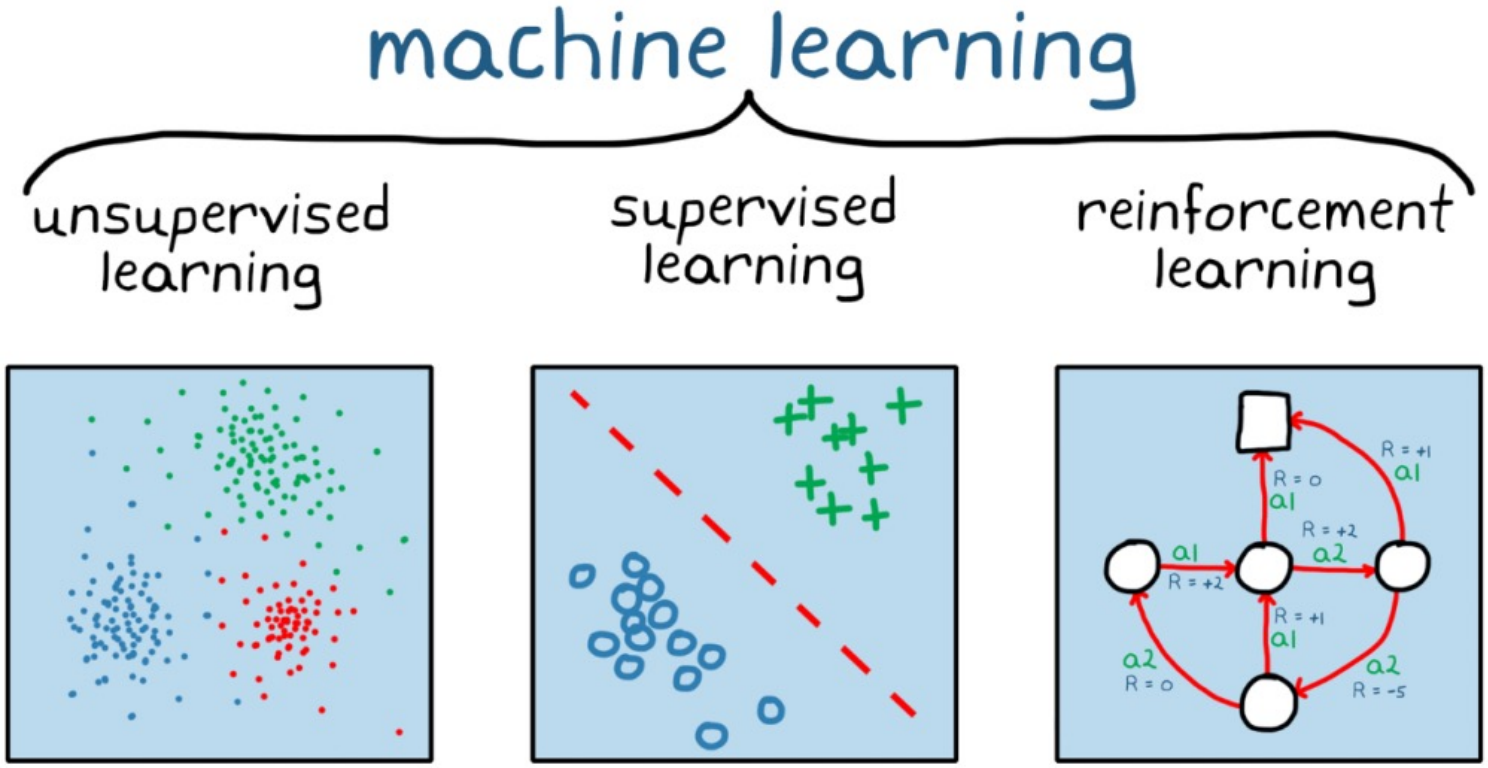
Transport and Logistic

- Route Optimization
- Traffic Prediction

What ML applications you know/being used in your country?

What ML applications you are interested in?

Types of machine learning



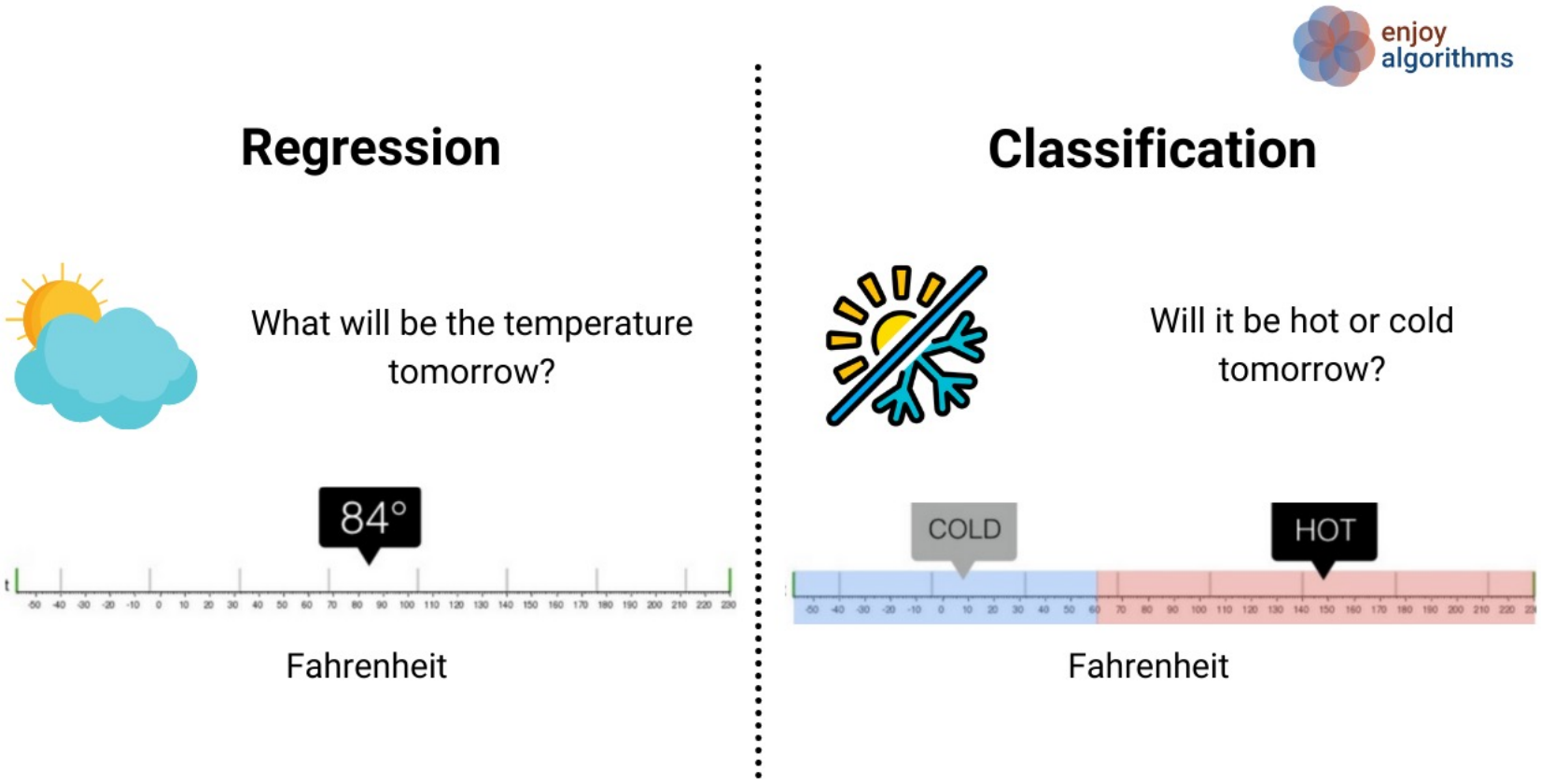
Unlabeled data
No feedback
Find hidden structure

Labeled data
Predict next value

Decision process
Reward system
Learn series of actions

Source: https://www.theaiops.com/machine-learning-compare-supervised-learning-vs-unsupervised-learning-vs-reinforcement-learning/

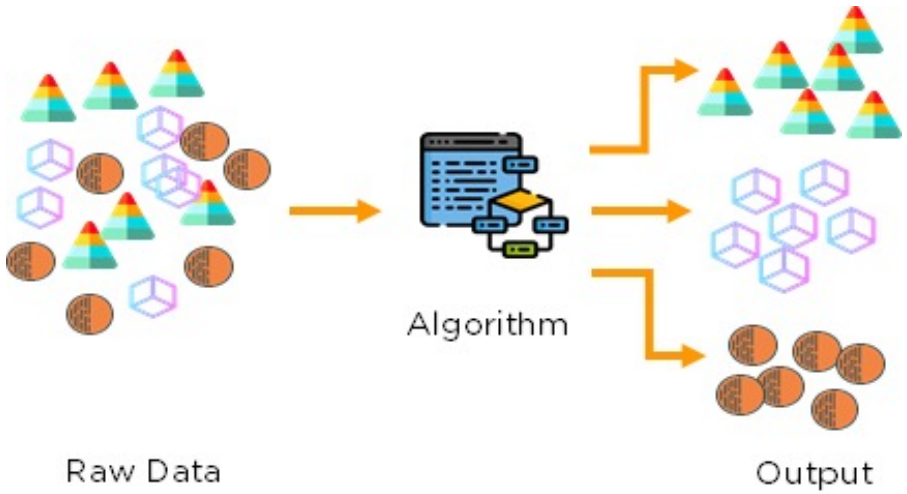
Types of Supervised Learning



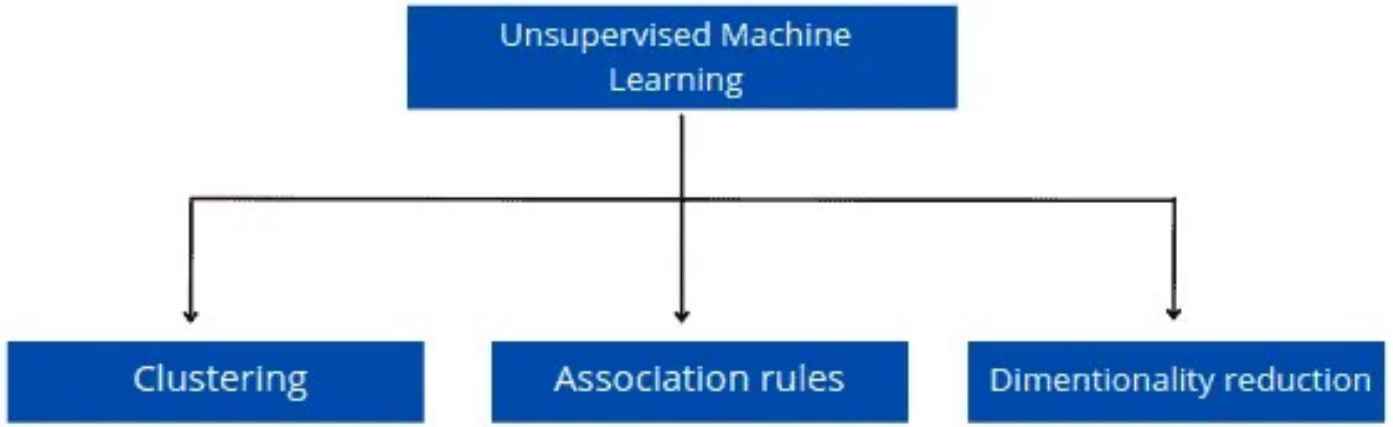
Binary/
Multi-class

Source: https://www.enjoyalgorithms.com/blogs/classification-and-regression-in-machine-learning

Types of Unsupervised Learning



Source: https://medium.com/analytics-vidhya/beginners-guide-to-unsupervised-learning-76a575c4e942



Source:https://hands-on.cloud/ml-unsupervised-learning-guide/

What type of ML problem do you think it is?

Scenario 1: A bank wants to predict whether a new applicant will default on a loan. They have data on previous applicants, including income, credit score, and past financial behavior.

Question: What type of machine learning problem is this, and why?

What type of ML problem do you think it is?

Scenario 2: An e-commerce company wants to group its customers into different segments to better tailor its marketing strategy. They have information on customer behavior, such as purchase frequency, spending habits, and browsing history, but they do not know how which customer belongs to which group.

Question: What type of machine learning problem is this, and why?

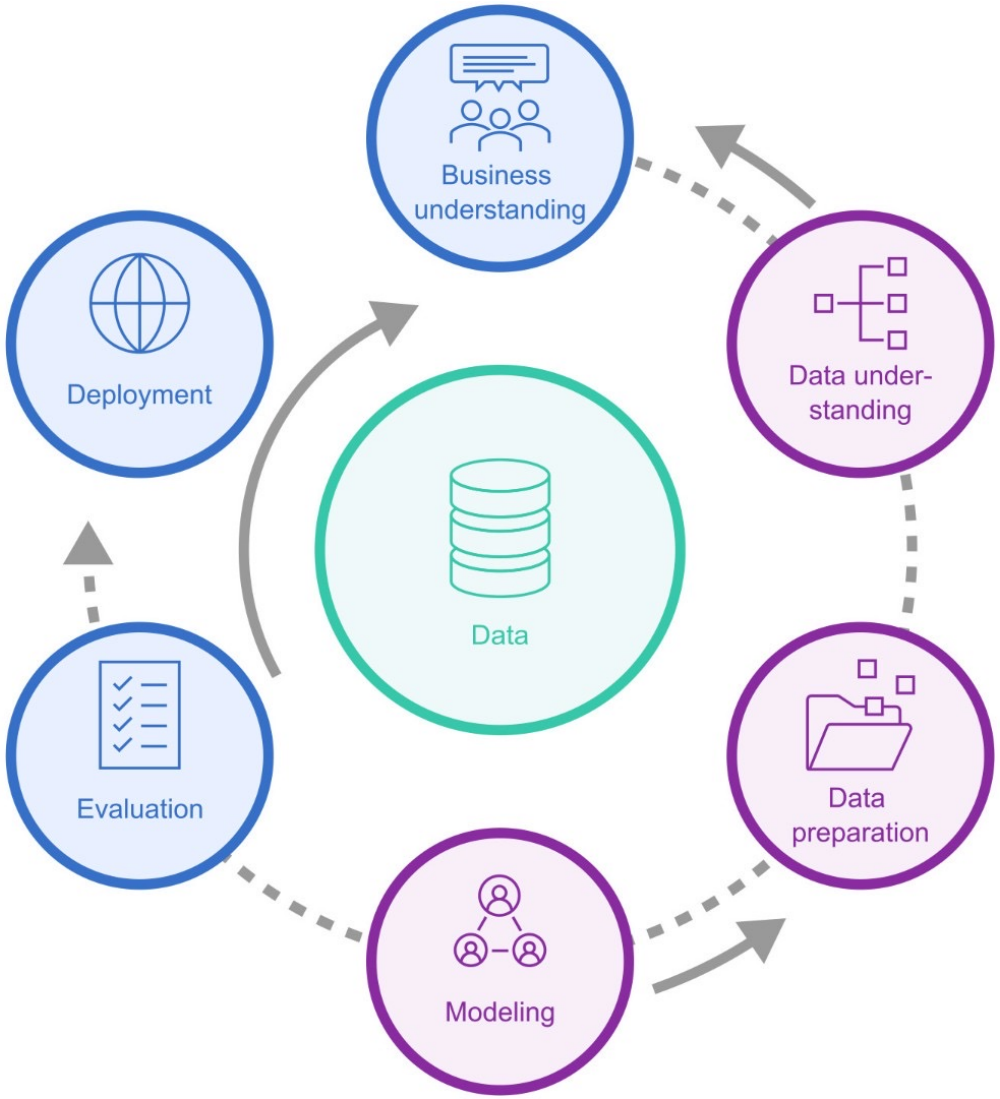
What type of ML problem do you think it is?

Scenario 3: Imagine a robotic vacuum cleaner that navigates around a house, cleaning floors and avoiding obstacles. The robot has to learn how to move efficiently, avoid furniture, and return to its charging station when its battery is low. The robot doesn't start with any prior knowledge about the house layout or where obstacles are located.

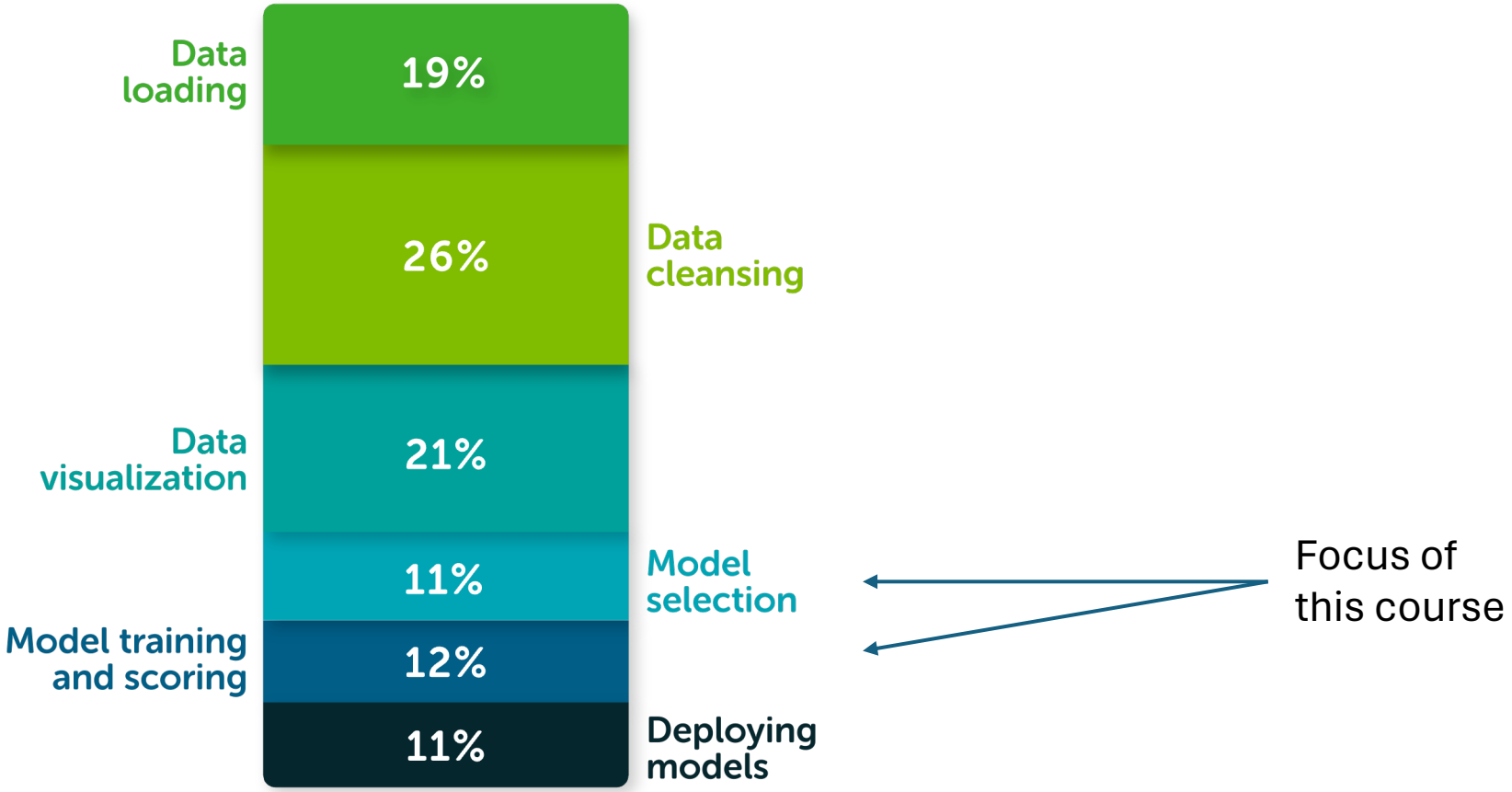
Question: What type of machine learning problem is this, and why?

Machine learning pipline

Cross-Industry Standard Process for Data Mining (CRISP-DM)



Time allocation for ML tasks



Source: https://www.anaconda.com/resources/whitepapers/state-of-data-science-2020

Components of a machine learning algorithm

- Input data
- Features
- Model
- Model output (prediction)
- Evaluation metrics

Some practical aspects

- Underfitting vs Overfitting
- Training, validation, and test dataset splitting
- Cross-validation
- Hyperparameters tuning
- Ensemble method

Sources to learn ML

- Scikit learn: https://scikit-learn.org/stable/
- Kaggle: https://www.kaggle.com/
- Machine Learning Mastery: https://machinelearningmastery.com/
- Books: https://github.com/josephmisiti/awesome-machine-learning/blob/master/books.md

Lab

Link:
https://github.com/luumsk/NSU_ML/blob/main/Labs/lab1.ipynb

References

- https://acropolium.com/blog/machine-learning-in-healthcare-use-cases-benefits-and-success-stories/
- https://www.enjoyalgorithms.com/blogs/classification-and-regression-in-machine-learning
- https://hands-on.cloud/ml-unsupervised-learning-guide/
- https://medium.com/analytics-vidhya/beginners-guide-to-unsupervised-learning-76a575c4e942
- https://www.erieri.com/salary/job/machine-learning-engineer/russian-federation
- https://www.tealhq.com/job-titles/machine-learning-scientist