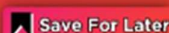# MACHINE LEARNING

DENIS PANJUTA @denis-panjuta

EXPLAINED IN A NUTSHELL

Save For Later

Computers use data to learn and make better predictions or decisions over time, applied in tasks like image recognition and recommendation systems.
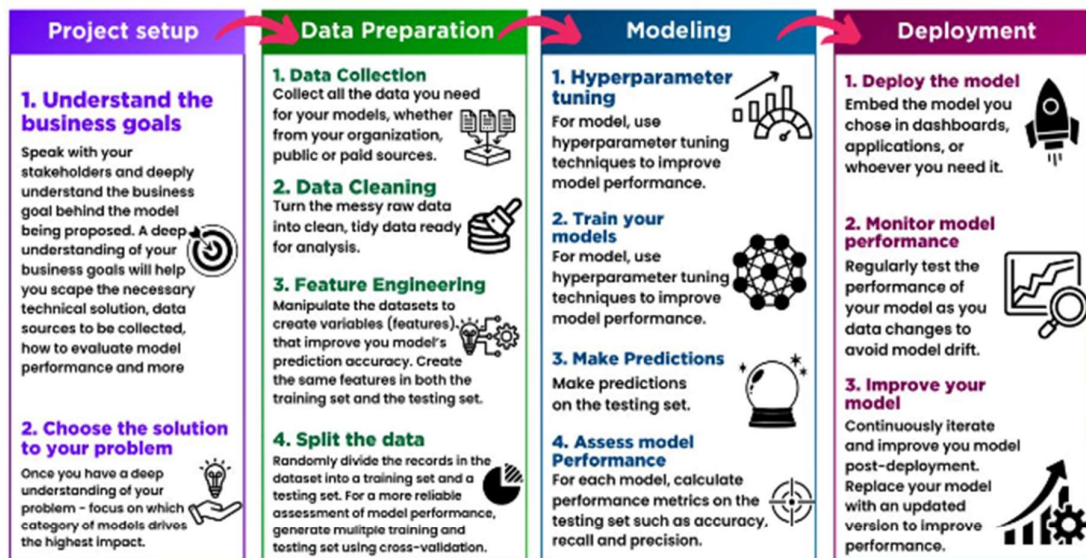
Machine Learning Job Roles

Machine Learning Engineer
- Model Maintenance
- Production
- APIs
- SQL
- Unit Testing
- Machine Learning Algorithms
- Deep learning Algorithms
- Predictive Web Apps
- R
- Python

BI Developers
- Business Actions
- Descriptive Dashboards
- Business Analysis
- Business Problems
- Shiny, Tableau, Excel, PowerBI, Qlik

NLP Engineer
- PDF, Word, Emails
- Video Processing
- Audio & Speech Processing
- Text Analysis
- Customers
- Communication
- Research
- Experiments
- Statistical Analysis

Data Scientist

3 Types of Machine Learning

Supervised Learning
- Labeled Data
- Direct Feedback
- Predict outcome / future

Unsupervised Learning
- No labels
- No feedback
- Find hidden structure in data

Reinforcement Learning
- Decision Process
- Reward system
- Learn series of actions

A BEGINNER'S GUIDE TO MACHINE LEARNING WORKFLOW

Project setup

1. Understand the business goals

Speak with your stakeholders and deeply understand the business goal behind the model being proposed. A deep understanding of your business goals will help you scape the necessary technical solution, data sources to be collected, how to evaluate model performance and more

2. Choose the solution to your problem

Once you have a deep understanding of your problem - focus on which category of models drives the highest impact.

Data Preparation

1. Data Collection
Collect all the data you need for your models, whether from your organization, public or paid sources.

2. Data Cleaning
Turn the messy raw data into clean, tidy data ready for analysis.

3. Feature Engineering
Manipulate the datasets to create variables (features) that improve you model's prediction accuracy. Create the same features in both the training set and the testing set.

4. Split the data
Randomly divide the records in the dataset into a training set and a testing set. For a more reliable assessment of model performance, generate multiple training and testing set using cross-validation.

Modeling

1. Hyperparameter tuning
For model, use hyperparameter tuning techniques to improve model performance.

2. Train your models
For model, use hyperparameter tuning techniques to improve model performance.

3. Make Predictions
Make predictions on the testing set.

4. Assess model Performance
For each model, calculate performance metrics on the testing set such as accuracy, recall and precision.

Deployment

1. Deploy the model
Embed the model you chose in dashboards, applications, or whoever you need it.

2. Monitor model performance
Regularly test the performance of your model as you data changes to avoid model drift.

3. Improve your model
Continuously iterate and improve you model post-deployment. Replace your model with an updated version to improve performance.

ML TERMS EXPLAINED

Algorithm
A set of instructions that tells a computer how to solve a specific problem step by step.

Supervised Learning
A type of machine learning where the computer learns from labeled examples provided by humans.

Label
The correct answer or output that the model is trying to learn and predict from the data.

Clustering
Grouping similar data points together without predefined labels, like organizing books by topic.

Data
Information collected and used by computers to learn, make decisions, and predict outcomes.

Unsupervised Learning
A type of machine learning where the computer finds patterns in data without any labeled examples.

Overfitting
When a model learns too much from the training data and performs poorly on new data.

Accuracy
A measure of how often the model's predictions are correct, shown as a percentage.

Model
A computer's representation or understanding of data used to make predictions or decisions.

Neural Network
A computer system designed to work like the human brain to recognize patterns and make decisions.

Classification
The process of sorting data into categories or groups, like sorting animals into types.

Bias
A systematic error in the model's predictions, often due to assumptions in the learning process.

Training
The process of teaching a computer by showing it many examples so it can learn patterns.

Feature
An individual measurable property or characteristic of the data used to make predictions.

Regression
A type of prediction where the goal is to predict a continuous number, like forecasting temperatures.

Validation
The process of checking how well the model works with new data, ensuring it can generalize to unseen data.