Machine Learning Concepts

1. Machine Learning Overview

An application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

- 🗭 Types of Machine Learning:
 - @ Supervised Learning:

The model learns from labeled data. It's like a student learning under the supervision of a teacher.

- Q Unsupervised Learning:

The model learns **from** unlabeled data. It's like learning **through** self-study **without** any guide.

- 🤝 Semi-supervised Learning:

The model learns from a mix of labeled and unlabeled data.

- 🏆 Reinforcement Learning:

The model learns **to** make decisions **by** performing certain actions **and** receiving rewards **or** penalties.

- M Bias-Variance Tradeoff:

It's about balancing the simplicity and accuracy of the model. A high-bias model could miss relevant relations between features and target outputs (underfitting), while a high-variance model could model the random noise in the training data (overfitting).

- • Overfitting and Underfitting:
 - 🍂 Overfitting:

When a model learns the detail and noise in the training data to the extent that it negatively impacts the model's performance on new data.

- 🌳 Underfitting:

When a model cannot capture the underlying trend of the data, leading to poor performance.

2. Machine Learning Algorithms

- 🥯 Linear Regression:

A linear approach to modeling **the** relationship between **a** dependent variable **and** one **or** more independent variables.

- 🌲 Decision Trees:

A type of supervised learning algorithm that is mostly used **for** classification problems.

- 🌠 Random Forests:

An ensemble learning method for classification, regression, and other tasks that operate by constructing a multitude of decision trees at training time.

- 📊 Logistic Regression:

A statistical **model** that uses a logistic function to **model** a **binary** dependent **variable**.

- 💠 Support Vector Machines (SVM):

A **set of** supervised learning methods used **for** classification, regression **and** outliers detection.

- \ K-nearest Neighbors (KNN):

A type of instance-based learning, **or** lazy learning, where the function is only approximated locally **and** all computation is deferred until function evaluation.

- 💫 K-Means Clustering:

A type of unsupervised learning used when you have unlabeled data (i.e., data without defined categories **or** groups).

- Principal Component Analysis (PCA):

A statistical procedure that uses **an** orthogonal transformation to convert **a** set **of** observations **of** possibly correlated variables **into a** set **of** values **of**

linearly uncorrelated variables called principal components.

- 🧩 Neural Networks:

A series **of** algorithms that endeavors to recognize underlying relationships **in a** set **of** data through **a** process that mimics **the** way **the** human brain operates.

3. Evaluation Metrics

- @ Accuracy:

The proportion of true results among the total number of cases examined.

- Precision:

The fraction of relevant instances among the retrieved instances.

- 🙏 Recall:

The fraction **of the** total amount **of** relevant instances **that** were actually retrieved.

- 🚀 F1 Score:

The harmonic mean of precision and recall.

- 1 Confusion Matrix:

A table used to **describe** the **performance** of a classification **model** on a **set** of data for which the true values are known.

- Area Under the ROC Curve (AUC-ROC):

A performance measurement ${f for}$ classification problem ${f at}$ various thresholds settings.

4. Feature Engineering

- 🧬 Feature Selection:

The process of reducing the number of input variables when developing a predictive model.

- 😂 Feature Extraction:

The process of transforming raw data into features that better represent the underlying problem to the predictive models.

- 🔄 Feature Encoding:

The process of converting categorical data to a form understandable by machine learning algorithms.

- Feature Scaling:

A method used to normalize the range of independent variables or features of data.

5.Deep Learning

- 🔡 Neural Networks:

Algorithms intended to mimic the human brain, that are capable of learning from data.

- * Activation Functions:

Functions that determine **the** output **of a** deep learning model, its accuracy, **and the** computational efficiency **of** training **a** model.

- 🏗 Convolutional Neural Networks (CNN):

A class **of** deep neural networks, most commonly applied **to** analyzing visual imagery.

- Necurrent Neural Networks (RNN):

A class of artificial neural networks where connections between nodes **form a** directed graph along **a** temporal sequence.

- ¾ Long Short Term Memory (LSTM):

A type of recurrent neural network capable of learning order dependence in sequence prediction problems.

6. Natural Language Processing

- 📖 Bag of Words:

A representation **of text** that describes **the** occurrence **of words within a** document.

- 🎭 Word Embedding:

The collective name **for** a set of language modeling and feature learning techniques in natural language processing.

- 🎢 N-grams:

A contiguous sequence of n items from a given sample of text or speech.

- // TF-IDF:

A numerical statistic **that is** intended **to** reflect how important a word **is to** a document **in** a collection **or** corpus.

- **C** Transformer Models:

An architecture **and model** family which uses self-attention mechanisms, known **for** the BERT, GPT **and** other **models**.

7. Reinforcement Learning

- 🎮 Q-Learning:

A model-free reinforcement learning algorithm **to** learn a policy telling an agent what action **to** take under what circumstances.

- 🎲 Monte Carlo Methods:

A broad class **of** computational algorithms **that** rely **on** repeated random sampling **to** obtain numerical results.

- 💫 Policy Gradients:

A type of reinforcement learning algorithms which are parameterized with neural network.

- 🧩 Multi-Armed Bandit:

A problem **in** which a fixed limited **set of** resources must be allocated **between** competing (alternative) choices **in** a way **that** maximizes their expected gain.

8. Advanced Topics

- 🌐 Federated Learning:

A machine learning approach where **the** training process is distributed **among** many users, but **the** model itself remains **on the host machine**.

- 🜒 Differential Privacy:

A system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset.

- ● Generative Adversarial Networks (GANs):

A class **of** artificial intelligence algorithms used **in** unsupervised machine learning, implemented **by a system of** two neural networks contesting **with each** other **in a** zero-sum game framework.

- Transfer Learning:

A machine learning method where a pre-trained model is used as the starting point for a different but related problem.

9. Probability and Statistics

- * Probability Distributions:

A **function** that describes the likelihood **of** obtaining the possible values that a random variable can assume.

- Normal Distribution:

A probability function **that** describes how **the** values **of** a variable are distributed. It **is** a symmetric distribution **where** most **of the** observations cluster **around the** central peak.

- M Binomial Distribution:

A probability distribution that describes **the** number **of** successes **in a** fixed number **of** binary experiments.

- Nypothesis Testing:

The process of using statistics to determine the probability that a specific hypothesis is true.

- Torrelation:

A statistical measure that expresses **the** extent to which two variables are linearly related.

10. Optimization Algorithms

- @ Gradient Descent:

An optimization algorithm used **to** minimize **some function by** iteratively moving **in** the direction **of** steepest descent **as** defined **by** the negative **of** the gradient.

- 🚵 Stochastic Gradient Descent:

A variation of the gradient descent algorithm that calculates the error and updates the model for each example in the training dataset.

- 🌄 Batch Gradient Descent:

A variation of the gradient descent algorithm which calculates the error for each example in the training dataset, but only updates the model after all training examples have been evaluated.

- 🌋 Mini-Batch Gradient Descent:

A compromise **between** batch gradient descent **and** stochastic gradient descent **where the** model **is** updated **after** a number **of** training examples.

11. Regularization Techniques

- X L1 Regularization:

A regularization technique which is used to shrink **the** coefficients **of the** model to zero, which can be used **for** feature selection.

- 📈 L2 Regularization:

A regularization technique which **is** used **to** prevent overfitting **by** discouraging high-valued coefficients **by** squaring their values.

- | Elastic Net:

A regularization technique which combines **both** L1 **and** L2 regularization. **It** can prevent model overfitting **and** perform feature **selection**.

- 📊 Early Stopping:

A form **of** regularization used **to** prevent overfitting when training a learner **with** an iterative **method**, **such as gradient descent**.

12. Dimensionality Reduction

- Principal Component Analysis (PCA):

A statistical procedure that uses **an** orthogonal transformation to convert **a** set **of** observations **of** possibly correlated variables **into a** set **of** values **of** linearly uncorrelated variables called principal components.

- ✓ Linear Discriminant Analysis (LDA):

A method used **in** statistics, pattern recognition, **and** machine learning to find **a** linear combination **of** features that characterizes **or** separates two **or** more classes **of** objects **or** events.

- 🗂 t-Distributed Stochastic Neighbor Embedding (t-SNE):

A machine learning algorithm **for** visualization based on stochastic neighbor embedding.

13. Ensemble Methods

- 🌳 Random Forest:

An ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees.

- 🎯 Boosting:

An ensemble meta-algorithm primarily used to reduce bias, and also variance in supervised learning.

- 📊 Bagging:

An ensemble meta-algorithm designed to improve **the** stability **and** accuracy **of** machine learning algorithms used **in** statistical classification **and** regression.

- 🌠 Stacking:

An ensemble learning technique that combines multiple classification or regression models via a meta-classifier or a meta-regressor.

14. Time Series Analysis

- ARIMA (AutoRegressive Integrated Moving Average):

A forecasting technique that projects the future values of a series based entirely on its own inertia.

- 🜊 Seasonality:

A characteristic **of** a time series **in** which **the** data experiences regular **and** predictable changes **that** recur **every** calendar year.

- ① Stationarity:

A property of a time series if its statistical properties such as mean, variance remain constant over time.

- 🔄 Autocorrelation:

A mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals.

15. Model Selection

- T Cross-validation:

A resampling procedure used to evaluate machine learning models on a limited data sample.

- 🚿 Bootstrap:

A powerful statistical method for estimating a quantity from a data sample.

- 🎯 Hyperparameter tuning:

The process of adjusting the configuration of the model to improve its performance.

- ① Grid search:

An approach to parameter tuning that will methodically build and evaluate a model for each combination of algorithm parameters specified in a grid.

16. Data Preprocessing

- 🔄 Data Cleaning:

The process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

- 🜊 Data Transformation:

The process **of** converting data from one format **or** structure **into** another format **or** structure.

- 🎈 Data Imputation:

The process of replacing missing data with substituted values.

- 🐧 Data Normalization:

The process of organizing data to minimize redundancy.

17. Other Machine Learning Algorithms

- 💫 Naive Bayes:

A classification technique based **on** Bayes' theorem **with** an assumption **of** independence among predictors.

- 🌲 Gradient Boosting:

A machine learning technique **for** regression **and** classification problems, which produces **a** prediction model **in the** form **of an** ensemble **of** weak prediction models.

- Nidge Regression:

A technique **for** analyzing multiple regression data **that** suffer **from** multicollinearity.

- 🜟 Lasso Regression:

A regression analysis method that performs both variable selection and regularization.

- 🎯 Elastic Net Regression:

A hybrid approach that blends both ridge regression and Lasso techniques.

18. Other Evaluation Metrics

- 📊 Log Loss:

The most important classification metric based on probabilities.

- ● Mean Absolute Error (MAE):

The mean of the absolute value of the errors.

- Mean Squared Error (MSE):

The mean of the squared errors.

- Root Mean Squared Error (RMSE):

The square root of the mean of the squared errors.

- R Squared (R2):

A statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable.

19. Recommended Study Resources

- Books:

- [The Hundred-Page Machine Learning Book](http://themlbook.com/): This book provides a great introduction to the concepts of machine learning.
- [Deep Learning](http://www.deeplearningbook.org/): A comprehensive book on deep learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
- [Pattern Recognition and Machine Learning](https://www.microsoft.com/en-us/research/people/cmbishop/#!prml-book): This book by Christopher Bishop provides an introduction to machine learning framed around probabilistic models.

- - Online Courses:

- [Coursera Machine Learning by Andrew Ng](https://www.coursera.org/learn/machine-learning): This course provides a broad introduction to machine learning, data mining, and statistical pattern recognition.
- [Udemy Machine Learning A-ZTM: Hands-On Python & R In Data Science](https://www.udemy.com/course/machinelearning/): This course provides a step-by-step guide to learn Machine Learning using Python and R.

- 🌐 Websites and Blogs:

- [Towards Data Science](https://towardsdatascience.com/): A Medium publication sharing concepts, ideas, and codes.
- [Machine Learning Mastery](https://machinelearningmastery.com/start-here/): This blog by Jason Brownlee provides practical advice and tutorials on machine learning.
- [KDnuggets](https://www.kdnuggets.com/): A leading site on AI, Analytics, Big Data, Data Mining, Data Science, and Machine Learning.

- 🎥 YouTube Channels:

- [3Blue1Brown](https://www.youtube.com/c/3blue1brown): This channel provides great visual explanations of mathematical concepts, including neural networks.
- [StatQuest with Josh Starmer](https://www.youtube.com/user/joshstarmer): This channel provides clear explanations of machine learning and statistics.
- [Siraj Raval](https://www.youtube.com/c/SirajRaval): This channel provides accessible educational videos on machine learning.