Understanding Machine Learning

Overview of Machine Learning

- **What is Machine Learning?**
  - Machine Learning (ML) enables computers to learn from data and improve over time without being explicitly programmed.
  - ML is widely used in real-life applications like intelligent assistants, recommendation systems, and medical diagnoses.
- **Machine Learning vs. Artificial Intelligence (AI)**
  - **AI**: Broad set of tools aimed at creating intelligent behavior in computers, including robotics, natural language processing, and machine learning.
  - **ML**: A subset of AI focused on algorithms that learn patterns from data to make predictions and inferences.
- **Key Concepts**
  - **Prediction**: Using historical data to predict future outcomes (e.g., weather forecast).
  - **Inference**: Drawing insights from data to understand causes or identify patterns (e.g., reasons behind weather changes).

Overview of Machine Learning

- **How Machine Learning Works**
  - ML relies on **data** and **algorithms** from statistics and computer science.
- **Learning from Data**:
  - ML algorithms analyze existing data to identify patterns and use this knowledge to make decisions on new, unseen data.
  - Example: Spam detection in emails—ML learns to identify spam from labeled email data and then can detect spam in new emails.
- **Role of Data Science**
  - Data science is about extracting insights and making sense of data.
  - ML is a tool within data science that helps make predictions and decisions from data.
- **Machine Learning Models**
  - **Model**: A statistical representation of a process (e.g., predicting traffic or classifying images).
  - The model is built from data, and new data can be input to make predictions or estimate outcomes (e.g., predicting traffic on a specific day).
  - The output can be a **probability** (e.g., likelihood that an image is of a cat).

ML Concepts

- **Reinforcement Learning**: Sequential decision-making based on rewards/penalties (e.g., robots, chess). Complex and less common.
- **Supervised Learning**: Uses labeled data (features + known outcomes) to train models for prediction (e.g., predicting heart disease from features like age, cholesterol). The model generalizes to new data.
- **Unsupervised Learning**: Finds patterns in unlabeled data (e.g., clustering, anomaly detection). Groups data based on similarity without predefined labels.
- **Training Data**: Data used to train models, with processing time varying based on dataset size.
- **Unlabeled Data**: Essential for unsupervised learning when labels are unavailable (e.g., self-driving car images).
- **Model Training**: The process of teaching a model using training data so it can make predictions or decisions.
- **Features**: Individual variables or pieces of data used to make predictions (e.g., age, cholesterol, etc. in heart disease prediction).
- **Labels**: Known outcomes in supervised learning that the model tries to predict (e.g., whether a patient has heart disease).
- **Generalization**: The model's ability to make accurate predictions on new, unseen data.
- **Applications**: Reinforcement learning is used in robotics and games, supervised learning is common in healthcare, finance, and marketing, while unsupervised learning is often used in customer segmentation, fraud detection, and anomaly detection.

ML Workflow

- **Extract Features**: Identify and prepare relevant features from the dataset (e.g., square footage, neighborhood). You may need to reformat or create new features based on the data.
- **Split Dataset**: Divide the dataset into two subsets: one for training the model (training dataset) and one for testing the model (test dataset). The test dataset ensures the model is evaluated on data it hasn’t seen before.
- **Train Model**: Input the training dataset into a chosen machine learning model. There are many models available (e.g., neural networks, logistic regression) depending on the complexity and use case.
- **Evaluate Model**: Test the model using the test dataset, often called "unseen data," to assess its performance. Common evaluation metrics include error rates or the percentage of correct predictions within a specified margin (e.g., 10% margin).
- **Tune Model (if necessary)**: If the model's performance isn't satisfactory, adjustments (or "tuning") can be made. This may involve tweaking model parameters or selecting different features. If performance remains poor, more data may be needed to improve results.

Supervised Learning

- Machine learning where the model is trained using labeled data to predict outcomes for new, unseen data.
- **Types of Supervised Learning**
  - **Classification**: Predicts discrete categories (e.g., yes/no, types of flowers, customer churn).
  - **Regression**: Predicts continuous values (e.g., stock price, temperature, real estate value).
- **Training Process**
  - **Data Splitting**: Split data into training (80%) and testing (20%) datasets.
  - **Model Training**: Use training data to build and train the model.
  - **Model Evaluation**: Test the model on unseen data and evaluate performance based on metrics like accuracy, error rate, etc.
- **Support Vector Machine (SVM)**
  - **Linear Classifier**: SVM separates categories using a straight line.
  - **Polynomial Classifier**: Allows curves for more complex, non-linear problems.
- **Choosing Between Classification & Regression**
  - **Classification**: Used when the target is categorical (e.g., yes/no).
  - **Regression**: Used when the target is continuous (e.g., price, temperature).
- **Model Performance & Tuning**
  - **Evaluation**: Assess model accuracy and make adjustments if needed.
  - **Tuning**: Improve performance by tweaking parameters or adding features.

Unsupervised Learning

- Machine learning where the model learns patterns from the dataset without a target column, identifying structures or groupings without pre-defined labels.
- **Applications**: Includes clustering, anomaly detection, and association.
- **Clustering**: Identifying groups of similar observations based on shared characteristics.
  - Examples: Grouping by species (dogs vs. cats), color (black, grey, white), or origin (Europe vs. Japan).
  - Models: K-Means (requires predefined clusters), DBSCAN (doesn't require predefined clusters).
- **Anomaly Detection**: Detecting outliers or data points that deviate significantly from the rest.
  - Examples: Identifying faulty devices, fraud detection, or unusual patient responses.
- **Association**: Finding relationships between events that occur together.
  - Examples: Market basket analysis (e.g., people buying jam also buying bread).

Evaluating Machine Learning Model Performance

- **Overfitting**: Overfitting occurs when a model performs well on training data but poorly on unseen testing data, meaning it has memorized the training set and fails to generalize well to new data.
- **Accuracy**: Accuracy is the ratio of correctly predicted observations to the total observations. However, it might not be the best metric in some cases, especially when dealing with imbalanced classes.
- **Confusion matrix**: The confusion matrix is a table used to evaluate the performance of a classification model. It includes
    - **True Positives (TP)**: Correctly predicted positive outcomes.
    - **False Positives (FP)**: Negative outcomes incorrectly predicted as positive.
    - **True Negatives (TN)**: Correctly predicted negative outcomes.
    - **False Negatives (FN)**: Positive outcomes incorrectly predicted as negative.

Evaluating Machine Learning Model Performance

- **Sensitivity**: Sensitivity (or Recall) measures the proportion of actual positives that are correctly identified. It focuses on minimizing false negatives.
- **Specificity**: Specificity measures the proportion of actual negatives that are correctly identified. It focuses on minimizing false positives.
- **Evaluating regression**: In regression, performance is evaluated based on the difference between the predicted values and the actual values. Metrics like Root Mean Squared Error (RMSE) or Mean Absolute Error (MAE) are commonly used to quantify this error.
- **Unsupervised learning evaluation**: In unsupervised learning, there are no target labels to compare predictions against, so the performance is evaluated based on how well the results meet the goals of the problem, such as grouping or detecting patterns.

Improving Model Performance

- **Evaluating Model Performance**: After training, if the model's performance isn't satisfactory, various techniques can be applied to enhance its effectiveness.
- **Dimensionality Reduction**: This technique involves reducing the number of features (or variables) in your dataset.
    - **Feature Selection**: Identifying and removing irrelevant or redundant features that don’t contribute to the model's predictive power.
    - **Feature Consolidation:** Reducing highly correlated features by merging them to retain important information while simplifying the model.
    - **Feature Transformation**: Combining multiple related features into a single feature to simplify the model without losing valuable information.
- **Hyperparameter Tuning**: Hyperparameters are settings that influence how a model is trained. Adjusting them can improve model performance.
    - **Adjustment**: Tuning hyperparameters such as the learning rate, regularization strength, or the number of layers in neural networks can help the model better learn the underlying patterns in the data.
    - **Optimization**: Structured methods, like grid search or random search, can systematically find the best set of hyperparameters to optimize the model’s performance.
- **Ensemble Methods**: Ensemble methods combine multiple models to make predictions more robust.
    - **Voting**: In classification tasks, the predictions of several models are combined, and the majority prediction is selected as the final output.
    - **Averaging**: In regression tasks, predictions from different models are averaged to produce a final prediction.

Deep Learning

- **What is Deep Learning?**
  Deep learning uses neural networks inspired by the human brain to solve complex problems, especially with unstructured data like text and images.
- **How it Works:**
  Neural networks process inputs (e.g., budget, advertising, star power) through multiple layers to predict outcomes (e.g., box office revenue).
- **Complex Networks:**
  As more data points are added, deep learning models become more complex, combining various factors like budget, advertising, and star power to make accurate predictions.
- **Training the Network:**
  The neural network learns the relationships between variables by analyzing training data, optimizing the connections between neurons.
- **When to Use Deep Learning:**
  Deep learning excels with large datasets and complex problems, particularly in areas like computer vision and natural language processing. For smaller datasets, traditional machine learning is more effective.

Computer Vision

- **What is Computer Vision?**
  Computer vision enables computers to interpret and understand images, essential for tasks like self-driving cars, which use cameras to detect objects, lanes, and traffic signs.
- **Image Data:**
  Digital images are made up of pixels, where each pixel has color and intensity values. For colored images, the RGB system is used, requiring three color channels (Red, Green, Blue).
- **Face Recognition:**
  In face recognition, images of people are input into a neural network, which learns to detect features like edges, eyes, and faces to ultimately recognize individuals.
- **Training the Neural Network:**
  By feeding labeled images (with the correct identity), the neural network learns what to compute for each neuron, gradually recognizing patterns like faces without needing manual intervention.
- **Applications:**
  Computer vision powers applications such as facial recognition, self-driving cars, medical image analysis (e.g., tumor detection), and even generating fake videos (e.g., deepfakes).

Natural Language Processing

- **What is NLP?:** NLP enables computers to understand and interpret human language, including tasks like identifying named entities (persons, locations) in text.
- **Bag of Words:** A simple NLP technique where word frequency in text is counted. While useful, it doesn't capture word context or meaning.
    - **N-grams:** By considering sequences of words (e.g., two-word sequences), n-grams improve upon bag of words, helping capture more context.
    - **Limitations of Bag of Words:** Word counts don't account for synonyms, such as different words meaning "blue" (e.g., sky-blue, aqua).
- **Word Embeddings:** A technique that groups similar words together, creating mathematical word representations. It enables intuitive relations, such as "King - man + woman = Queen."
- **Language Translation:** NLP techniques like word embeddings help translate text from one language to another, as seen in machine translation applications.
- **Applications of NLP:** NLP powers technologies such as Google Translate, chatbots, voice assistants (Siri, Alexa), and sentiment analysis tools.
- **Why Deep Learning for NLP?** Deep learning is preferred for text and image data because it doesn't need manual feature engineering and performs better as the amount of data increases.

Limitations of ML

- **Data Quality:** "Garbage in, garbage out" — The quality of the input data directly affects the quality of machine learning outputs. Poor data leads to inaccurate or biased results.
- **Real-world Examples of Data Issues:**
  - Amazon's AI recruiting system preferred male candidates due to biased training data from past hiring trends.
  - Microsoft’s chatbot Tay, corrupted by internet trolls, demonstrated the dangers of machine learning models learning from harmful inputs.
- **Importance of Critical Data Analysis:** Always be critical of the model's outputs. High-quality data is essential, including analysis of data characteristics, relevance, and domain expertise.
- **Quality Assurance:** To ensure data quality, it's important to assess the data source, distribution, and outliers, with transparent and repeatable processes.
- **Explainability:**
  - Machine learning models can be "black boxes," making it hard to understand their decision-making process.
  - Transparency is crucial for trust, regulatory compliance, and bias detection.
- **Explainable AI:** AI models must sometimes explain their reasoning to be useful in real-world applications. Explainable AI methods aim to clarify the factors influencing predictions.
- **Example of Explainable AI:** A hospital using a machine learning model for diabetes prediction can identify key features, like blood pressure, that impact predictions, providing valuable insights.
- **Inexplicable AI in Deep Learning:** For tasks like handwritten letter recognition, deep learning works well even without explainability, as accuracy is the primary goal.