DAVE3625 & DIKU 004 Artificial Intelligence Study Guide - Google Docs

Introduction to Artificial Intelligence (Oslomet - storbyuniversitetet)

DAVE3625 & DIKU 004: Artificial Intelligence Study Guide

Quiz

Questions

Answer each of the following questions in 2-3 sentences, using only the information provided in the source materials.

1. What is Machine Learning and what is its relationship to Artificial Intelligence (AI)?
2. Describe the core concept of Reinforcement Learning, including its feedback mechanism.
3. Explain the difference between Strong/General AI and Weak/Narrow AI. Which type is more relevant for current business applications?
4. What are the two biggest drivers behind the recent progress in deep learning and neural networks?
5. According to the materials, what is the primary purpose of development (dev) and test sets in a machine learning project?
6. Define "Core Business Data" and explain why it is a prime candidate for AI applications.
7. What are Generative Adversarial Networks (GANs) and what two components do they consist of?
8. Explain the concept of "data mismatch" and why it is important to diagnose.
9. What are the two main types of tasks in unsupervised learning? Provide a brief description of each.
10. Describe the concept of a "cost function" in machine learning and the purpose of the Gradient Descent algorithm in relation to it.

Answer Key

1. Machine learning is an application of AI that gives systems the ability to automatically learn and improve from experience without being explicitly programmed. It is a subset of the broader field of AI, focusing on techniques that allow machines to learn from data rather than being programmed with explicit rules.
2. Reinforcement Learning is a technique that enables an agent to learn in an interactive environment through trial and error. The learning process is guided by feedback; the

model is penalized for wrong predictions and rewarded for correct ones, based on its own actions and experiences.

3. Strong/General AI is occupied with the greater goal of reproducing and understanding intelligence on a human level. Weak/Narrow AI are systems able to perceive a narrow segment of the world and automate decisions, which is more relevant for applied research and business today because it can deliver more immediate value.
4. The two biggest drivers of recent progress are data availability and computational scale. People's digital activities generate vast amounts of data, and we now have the computational power to train neural networks large enough to take advantage of these huge datasets.
5. The purpose of dev and test sets is to direct a team toward the most important changes to make to a machine learning system. They allow the team to quickly evaluate ideas, tune parameters, and see how well an algorithm is performing on data that reflects what it will encounter in the future.
6. "Core Business Data" is the data closest to the value proposition of a business, which has a direct impact on the organization’s performance. Because any improvement in performance on this data is almost guaranteed to make a dramatic impact, and it's easy to attach monetary value to it, it is the prime candidate for AI applications.
7. Generative Adversarial Networks (GANs) are an important subset of Generative AI used to create realistic data like images, music, and text. They consist of two competing neural networks: a generator that creates data and a discriminator that evaluates it, aiming to distinguish the generated data from real data.
8. Data mismatch occurs when the training set data is a poor match for the dev/test set data, causing a model that generalizes well on the training distribution to perform poorly on the target distribution. Diagnosing it is critical to understand if poor performance is due to the model's inability to learn (bias/variance) or because the training data is not representative of the real-world problem.
9. The two main tasks in unsupervised learning are clustering and association. Clustering is used to discover natural groupings within the data. Association is used to discover rules that describe the data, such as finding frequent itemsets in transaction data.
10. A cost function is a measure used in machine learning to estimate how badly a model is performing, or how wrong it is in its ability to estimate the relationship between inputs and outputs. Gradient Descent is an optimization algorithm used to iteratively tweak the model's parameters to minimize this cost function, thereby reducing errors.

Essay Questions

The following questions are designed for longer, essay-style answers. You should synthesize information from across the provided source materials to formulate your response. No answers are provided.

1. Andrew Ng's text "Machine Learning Yearning" emphasizes the importance of a rapid, iterative approach to building ML systems, starting with a basic system and using error analysis to guide improvements. Using examples from the source texts (such as the cat detector or spam filter), explain this iterative process. Discuss the specific steps of error analysis and how it helps prioritize tasks.
2. Compare and contrast the three main types of machine learning: Supervised, Unsupervised, and Reinforcement Learning. For each type, describe its core purpose, the kind of data it requires (labeled vs. unlabeled), a primary challenge associated with it, and provide at least one specific algorithm example mentioned in the texts.
3. The source materials describe both a "Typical M.L Flow" and the significant time AI programmers spend on data-related tasks. Synthesize these sources to create a comprehensive overview of the data preparation and management lifecycle in an AI project. Your answer should cover topics like data collection, labeling, cleaning, feature engineering, and potential data pitfalls.
4. Discuss the evolution of Generative AI, from early concepts of neural networks to the development of GANs, and finally to the Transformer architecture that enabled Large Language Models (LLMs). Explain the key breakthrough that each of these developments represented and how they built upon one another.
5. The texts highlight a significant gap between the "hype of A.I" and the "reality of A.I." Using the examples of driverless cars, AI in healthcare, and business adoption, analyze the primary limitations and challenges that have slowed the real-world implementation of advanced AI systems.

Glossary

Term	Definition
Apriori Algorithm	An unsupervised association algorithm used for mining frequent itemsets and devising association rules, often used to power "You may also like" features.
Artificial Intelligence (AI)	A field encompassing any trait exhibited by machines considered similar to human intelligence. A machine has AI if it can interpret data, learn from it, and use that knowledge to adapt and achieve specific goals.

Association	A type of unsupervised learning that aims to discover rules describing the data, such as relationships between items in a dataset.
Avoidable Bias	The difference between the training error and the optimal (or human-level) error rate. It reflects how much worse an algorithm performs on the training set than the "optimal classifier."
Bias	An error source in machine learning. Informally, it is the algorithm's error rate on the training set, indicating how well the model fits the training data.
Blackbox Dev Set	A subset of the development set used for automatic evaluation (measuring error rates, tuning hyperparameters) that the developer should avoid looking at manually.
Classification	A supervised learning task that divides data into two or more well-defined classes or categories, such as spam detection (spam or normal).
Clustering	An unsupervised learning task that aims to discover natural groupings in unlabeled data.
Context Window	In LLMs, the maximum span of text (measured in tokens) that a model can consider at any given time when processing input and generating responses.
Core Business Data	The data closest to the value proposition of a business, describing events and patterns that have a direct impact on the organization’s performance.

Cost Function	A measure of how wrong a machine learning model is in terms of its ability to estimate the relationship between inputs (X) and outputs (y). It is used to estimate how badly a model is performing.
Data Mismatch	A problem that occurs when the training data distribution is a poor match for the dev/test set distribution, causing poor performance on the target task.
Data Mining	The process of discovering patterns in large data sets, involving methods at the intersection of machine learning, statistics, and database systems.
Data Wrangling	The process of transforming and mapping data from a "raw" data form into another format to make it more appropriate and valuable for downstream purposes like analytics.
DBSCAN	(Density-Based Spatial Clustering of Applications with Noise) A density-based clustering algorithm that separates regions by areas of low-density to detect outliers between high-density clusters.
Deep Learning (DL)	A subset of machine learning that uses neural networks, often with many layers. All main kinds of machine learning can leverage deep learning.
Development (Dev) Set	A dataset used to tune parameters, select features, and make other decisions regarding the learning algorithm. Also called a hold-out cross validation set.
End-to-End Learning	An approach where a single learning algorithm, typically a neural network, learns to go directly from the input to the desired output, replacing a multi-step pipeline.

Error Analysis	The process of examining dev set examples that an algorithm misclassified to understand the underlying causes of the errors, helping to prioritize projects.
Eyeball Dev Set	A subset of the development set that a developer manually examines to perform error analysis and gain intuition.
Feature Engineering	The process of constructing new features from existing data to train a machine learning model. Also known as feature creation.
Foundation Model	A category for a very large, versatile model (like a big toolbox) that can be adapted to do many different tasks. LLMs are a type of foundation model.
Fuzzy Logic	A way of dealing with uncertainty in Symbolic AI by using values between 0 and 1, rather than strict binary true/false (0 or 1).
Generative Adversarial Network (GAN)	A type of generative AI model consisting of two competing neural networks—a generator and a discriminator—that leads to the generation of highly realistic data.
Generative AI	A branch of AI and a subset of traditional machine learning that creates new, novel content in forms such as text, images, or audio.
Gradient Descent	An optimization algorithm used to tweak a model's parameters iteratively to minimize the cost function by learning the direction that reduces errors.

K-Means Clustering	A distance-based unsupervised clustering algorithm that calculates distances between data points to assign them to a predefined number (K) of clusters.
Knowledge Base	In Symbolic AI, the collection of all symbols and propositions that are considered true, representing the entire world for the software.
Large Language Model (LLM)	An advanced AI model, built on a transformer architecture, trained on vast datasets to understand and generate human language.
Learning	In the context of machines, the creation of models from data. It is defined as changes in behavior as a result of experience.
Logistic Regression	A supervised learning algorithm used for binary classification problems (e.g., yes/no, pass/fail) that uses a sigmoid function to transform any value into a probability range of 0 to 1.
Machine Learning (ML)	An application of AI that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.
Naive Bayes Classifier	A supervised classification algorithm that assumes the presence of a particular feature in a class is unrelated to the presence of any other feature.
Narrow AI (Weak AI)	AI systems able to perceive a narrow segment of the world, perform operations, and deploy decisions in response. It is essentially the automation of decisions.

Neural Network	A type of machine learning model inspired by the human brain's structure, designed to recognize patterns and make predictions.
Optimizing Metric	A single-number evaluation metric that a team tries to maximize or minimize, such as accuracy.
Probability	The likelihood of something happening, calculated as the number of ways it can happen divided by the total number of outcomes.
Recommender Systems	A class of machine learning algorithms that offer relevant suggestions to users by predicting the most likely product they will be interested in.
Regression	A supervised learning task used to predict a continuous value, such as predicting the price of a house.
Reinforcement Learning	A type of machine learning where an agent learns in an interactive environment by trial and error, using feedback (rewards and penalties) from its own actions.
Satisficing Metric	An evaluation metric where performance only needs to be "good enough" (i.e., meet a certain threshold), rather than being maximized.
Statistics	A branch of mathematics dealing with data collection, organization, analysis, interpretation, and presentation.
Strong AI (General AI)	AI concerned with the goal of reproducing and understanding intelligence to a degree comparable to humans across many different fields.

Supervised Learning	A family of machine learning algorithms that allows computers to learn how to map inputs (features) to outputs (labels) given enough labeled examples.
Support Vector Machine (SVM)	A supervised learning algorithm that can be used for classification and regression. In classification, it works by finding a hyperplane that creates a clear separation between data points of different classes.
Symbolic AI (GOFAI)	"Good Old-Fashioned AI" based on high-level symbolic representation of problems, logic, and search. It separates logic from data and uses programmed rules.
Test Set	A dataset used to evaluate the final performance of an algorithm, but not to make any decisions regarding the algorithm's design or parameters.
Token	A basic unit of text used by LLMs, which can be a word, part of a word, or punctuation.
Training Set	The dataset on which a learning algorithm is run to learn patterns and create a model.
Transformer Model	A neural network architecture, introduced in 2017, that relies on an attention mechanism. It revolutionized natural language processing and is the foundation for most modern LLMs.
Unsupervised Learning	A type of machine learning that looks for undetected patterns and learns to identify groups in unlabeled data with minimal human supervision.

Variance	An error source in machine learning. Informally, it measures how much worse an algorithm performs on the dev/test set than on the training set, indicating how well the model generalizes.