Science

Course:
INFO-6145 Data Science and Machine Learning

Revised by:
Mohammad Noorchenarboo

November 20, 2024

Contents

Current Section

Types of Reasoning

Deduction

- Starts with general rules to reach a specific conclusion.
- The conclusion must be true if the rules are true.

Example: "All humans are mortal; David is human; therefore, David is mortal."

Induction

- Looks at specific examples to make a general rule.
- The conclusion might be true but is not guaranteed.

Example: "If you see three white golf balls, you might think all the golf balls in the bag are white."

Abduction

- Picks the most likely explanation based on what is known.
- Often used when there is no clear proof.

Example: "A detective solves a crime by connecting clues to figure out who did it."

Rationalism vs. Empiricism

Rationalism

- Believes senses can be misleading, so trust logic and reasoning.
- Common in areas like math, coding, and designing theories.

Math: “Proving a formula without experiments, using logic alone.”

Empiricism

- Believes we learn by observing and experimenting.
- Common in areas like medicine, science experiments, and fixing software bugs.

Medicine: “A doctor diagnoses an illness by observing symptoms and running tests.”

Hume's Problem of Induction

The Problem:

- Induction is when we make general conclusions based on specific observations.
- **Hume's Question:** How do we know future events will always follow the same patterns as past events?

Example

A machine learning model predicts house prices based on historical data. But how do we ensure the model's predictions remain accurate if the housing market undergoes unforeseen changes? This assumption of consistency is central to Hume's problem.

Hume's Problem of Induction

Why It Matters

- Science and machine learning rely heavily on patterns from the past to make predictions.
- However, those predictions might fail if:
  - The patterns change.
  - We encounter new situations the model has not seen before.

Example Questions

- "When will you trust a COVID vaccine?" – Does the success of past vaccines guarantee the effectiveness of a new one?
- "When will you feel safe riding in a self-driving car?" – Can a car handle every possible situation it hasn't encountered before?

Current Section

What is Science?

Common View: Science is often seen as a step-by-step process:

1. **Observe:** Gather information about the phenomenon or system being studied.
2. **Hypothesize Based on Observation:** Propose a testable and falsifiable explanation for what you observed.
3. **Predict:** Make specific predictions about what should happen if the hypothesis is correct.
4. **Test:** Conduct experiments or collect data to evaluate the accuracy of your predictions.
5. **Analyze Results:** Review findings to determine if they support or refute the hypothesis. Revise the hypothesis and repeat the process if necessary.

What is Science?

Examples

- **Physics:** Observing that apples fall from trees and hypothesizing that gravity pulls them downward.
- **Medicine:** Testing the efficacy of a new drug by administering it to a controlled group and analyzing outcomes.
- **Machine Learning:** Building a recommendation system, hypothesizing that user history predicts preferences, testing the model, and refining it using evaluation metrics.
- **Astronomy:** Observing the orbits of planets, hypothesizing the influence of gravity, and validating with mathematical models.
- **Ecology:** Studying deforestation rates, hypothesizing their impact on climate change, and analyzing global temperature data.

What is Science?

Scientific Reality:

- **No Single Method:** Contrary to popular belief, there is no universally fixed “scientific method.” Scientists adapt their approaches based on the problem and context. For example, observational sciences like astronomy rely on data collection, while experimental sciences like chemistry focus on controlled experiments.
- **Non-linear and Flexible Processes:** Science is not always a step-by-step process. Discoveries often happen through iterative cycles, unexpected findings, or creative leaps. For instance, breakthroughs like penicillin were discovered accidentally and followed by systematic investigation.

What is Science?

- **Key Principles:** Science is guided by fundamental principles rather than rigid rules:
  - **Empirical Evidence:** Scientific conclusions must be based on measurable and observable data. For example, a new drug must demonstrate its efficacy through clinical trials rather than mere speculation.
  - **Theories Evolve with New Evidence:** Scientific theories are not absolute truths but models of understanding. They are revised or replaced when new evidence contradicts them. For instance, Newtonian mechanics was expanded by Einstein's theory of relativity to account for phenomena at very high speeds and gravitational fields.
  - **Extraordinary Claims Require Extraordinary Evidence:** Bold scientific claims, such as the possibility of faster-than-light travel, must be backed by exceptional proof. This principle emphasizes that the more a claim challenges existing understanding, the stronger the evidence required to support it.

Hypotheses and Theories

Hypothesis

- A hypothesis is a specific, testable statement about the world.
- Example: “A high-sugar diet increases the risk of developing Type 2 diabetes.” This can be tested by studying populations with varying sugar intake and measuring diabetes incidence.
- Hypotheses cannot be “proven” correct but can be:
  - **Corroborated:** Supported by evidence. For example, a study finds that individuals with high sugar consumption are consistently at greater risk of developing Type 2 diabetes compared to those with low sugar intake.
  - **Refuted:** Disproved by counterexamples. For example, finding a population with a high-sugar diet and a low prevalence of Type 2 diabetes.
- Inspired by Newton’s principle: “Whatever is true of what we’ve seen is true of the universe.” This principle highlights the importance of observing patterns to form hypotheses while acknowledging that exceptions may exist.

Hypotheses and Theories

Theory

- A theory is a well-supported explanation of the natural world.
- Developed when a hypothesis withstands extensive evidence and peer review.
- Examples:
  - **Biology: Germ Theory of Disease:**
    - Hypothesis: Microorganisms cause infections.
    - Evidence: Repeated experiments (e.g., Pasteur's work) proved consistent with the hypothesis.
    - Theory: Germ theory explains why antiseptic practices reduce infections.
  - **Physics: Theory of Gravity:**
    - Hypothesis: Objects with mass attract each other.
    - Evidence: Observations of planetary motion and falling objects match predictions.
    - Theory: Newton's theory of gravity explains these interactions (later refined by Einstein's general relativity).

Hypotheses and Theories

Theory

- **Earth Science: Plate Tectonics Theory:**
  - Hypothesis: The Earth's crust is divided into plates that move over time.
  - Evidence: Matching fossils across continents, patterns of earthquakes, and seafloor spreading.
  - Theory: Plate tectonics explains phenomena such as earthquakes, volcanic activity, and continental drift.

- A valid theory must:
  1. Predict real-world phenomena (e.g., germ theory predicts the spread of infection; gravity predicts planetary motion).
  2. Explain why predictions work (e.g., microorganisms multiply; gravitational force depends on mass and distance).
  3. Survive peer review by experts.

Life of a Theory

- Theories are refined and adjusted over time to fit new evidence.
- Older theories are replaced by newer ones when they no longer align with observations.

Standards for Theories

- **Requirements for Theories:**
  - Must predict real-world phenomena and anticipate new evidence.
  - Should explain why predictions are expected to work.
  - Must survive peer review, enabling other scientists to replicate data fits and predictions.

Example

- Theory of evolution predicts patterns in the fossil record and genetic similarities between species.
- Peer-reviewed studies consistently validate these predictions.

The Provisional Nature of Theories

Key Idea

- Theories can never be “proven,” only corroborated by evidence.
- They remain **provisional** even with extensive supporting evidence.

Example: Germ Theory of Disease

- Germ theory is supported by countless experiments.
- Even so, it is still provisional, leaving room for refinement with new evidence.

Utility of Theories

- Despite being provisional, theories provide reliable explanations for observed phenomena.
- Example: Newtonian physics remains useful for engineering and navigation despite being superseded by general relativity.

Scientific Refinement of Theories

- Theories undergo years of adjustment to fit new evidence.
- Some adjustments succeed only partially, leading to the creation of better theories.
- Example: The transition from the Ptolemaic geocentric model to the Copernican heliocentric model.

Refinement Example

- Newtonian mechanics worked well for most physical systems.
- When applied to very high speeds or strong gravitational fields, it failed to match observations.
- Einstein's theory of relativity refined these principles, explaining phenomena like time dilation and gravitational lensing.

Scientific Revolution

Revolutionary Shift in Theories

- When new theories better fit evidence, they replace older ones.
- Example:
  - Einstein's general relativity replaced Newtonian gravity.
  - Relativity explained phenomena Newtonian gravity could not, such as the bending of light by massive objects.

Importance of Evidence

- The strength of a new theory depends on its ability to explain existing data and predict new phenomena.
- Example: Quantum mechanics replaced classical mechanics at microscopic scales due to its greater explanatory power.

Feynman's Insight on Science

Core Principle

- Richard Feynman emphasized that **theories must agree with experimental results.**
- "If a theory disagrees with experiment, it is wrong," no matter how elegant it seems.

Discarding Incorrect Theories

- The essence of science is the ability to reject theories when they do not align with evidence.
- Example:
  - **Early AI Models:** Rule-based systems were initially thought to be sufficient for building intelligent systems. However, they failed to handle unstructured data like images and natural language.
  - **Modern Understanding:** These systems were replaced by machine learning models, such as neural networks, which learn patterns directly from data.

Lifecycle of a Theory

- Theories evolve as they are tested and refined.
- Over time, evidence accumulates, leading to further refinement or replacement.

Example: Plate Tectonics

- Initial hypothesis: Continents drift.
- Evidence: Matching fossils across continents, seafloor spreading.
- Result: Developed into the modern theory of plate tectonics, explaining earthquakes, volcanic activity, and mountain formation.

Generational Refinement

- New generations of scientists improve theories as technology and knowledge advance.
- Example: Advances in molecular biology have refined the theory of evolution with detailed genetic evidence.

Machine Learning and Science

Parallels with Scientific Processes

- Machine learning mirrors the scientific process:
  1. **Generating hypotheses:** Algorithms identify patterns (e.g., “These features predict housing prices”).
  2. **Testing hypotheses:** Models are tested on unseen data to validate predictions.
  3. **Refining or discarding hypotheses:** Poor-performing models are updated or replaced.
- Example:
  - A model predicts stock prices based on historical data.
  - It is tested on new data and refined to improve accuracy.
- Reflects the iterative nature of science: Hypothesize, test, revise, and repeat.

Current Section

References

- **The Scientific Attitude** by Lee McIntyre.
- **The Master Algorithm** by Pedro Domingos.