Northwestern Kellogg

Northwestern ENGINEERING

MBAi 448 | Winter 2026

Deep Learning

Today

- Assignment 3 / week 4 takeaways	*10 minutes*
- How deep learning helps	*10 minutes*
- Is deep learning the tool for that job?	*20 minutes*
- Quick break	*5 minutes*
- Deep learning case studies	*25 minutes*
- Assignment walkthrough	*20 minutes*

Assignment 3 / week 4 takeaways

Deep Learning

Week 4 takeaways

1. Artificial Neurons & Neural Network Foundations

Core idea: How neural networks are built from simple computational units.

- Artificial neuron = weighted inputs + bias → activation function
- Linear combination wTx+bw^T x + bwTx+b followed by nonlinearity
- Biological neuron analogy used to build intuition (early slides)
- Distinction between **shallow** (1 hidden layer) vs **deep** (2+ hidden layers) networks

👉 This sets the mental model for everything that follows.

2. Activation Functions & Their Learning Tradeoffs

Core idea: Why non-linearities matter and how activation choice affects learning.

- Common activations: **Sigmoid, Tanh, ReLU, Leaky ReLU, Softmax, Swish**
- Visualization of activation curves and their “flat” vs “steep” regions
- Discussion of vanishing gradients and exploding signals through depth
- Output-layer activations tied to task type (binary vs multi-class)

👉 A practical lens on *why ReLU dominates hidden layers* and *why output activations are task-specific*.

Week 4 takeaways

3. Training Dynamics: Loss, Backpropagation, and Optimization

Core idea: How neural networks actually learn.

- Training loss = difference between predictions and ground truth
- Backpropagation as gradient descent implemented via the **chain rule**
- Key training terms: **batch, iteration, epoch**
- Optimization framed as minimizing a cost function (MSE shown explicitly)

👉 This explains *how weights change* and *why depth makes training harder*.

4. Representation Learning & Feature Hierarchies

Core idea: Deep learning learns features automatically — layer by layer.

- Early layers learn low-level patterns (edges, textures)
- Middle layers learn shapes and parts
- Late layers learn semantic concepts (objects, decisions)
- Visual examples of how networks “see” images across layers

👉 This is the conceptual heart of deep learning vs traditional feature engineering.

Week 4 takeaways

5. Major Deep Learning Architectures

Core idea: Different architectures specialize in different data types.

- **CNNs** for spatial data (images, invariance to translation/scale/rotation)
- **RNNs / LSTMs** for sequences and time dependencies
- **Transformers** for attention-based sequence modeling (modern foundation models)

👉 Helps answer "*why this model for this problem?*"

6. Pre-Training, Transfer Learning, and Fine-Tuning

Core idea: Why modern AI rarely trains from scratch.

- Pre-training on massive datasets (expensive, slow, compute-heavy)
- Transfer learning: reuse learned representations
- Fine-tuning on smaller, domain-specific datasets
- Concrete contrast between base models and task-specific models

👉 This bridges classical DL with today's foundation models and LLMs.

Week 4 takeaways

7. Connectionist vs Symbolic AI (and Hybrid Systems)

Core idea: Where neural networks are strong — and where they are dangerous alone.

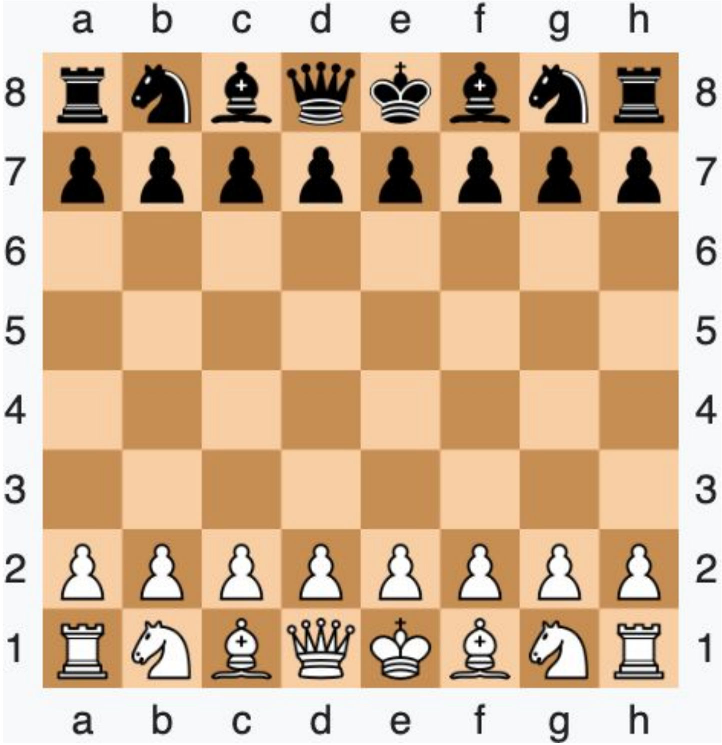
- **Connectionist AI**: pattern recognition, learning from data, black-box behavior
- **Symbolic AI**: rules, logic, transparency, explicit reasoning
- Activities focused on **risk, trust, refusal, and escalation**
- Emphasis on **hybrid systems** when wrong answers are costly

👉 This is the governance + decision-making lens layered on top of the tech.

How deep learning helps

Deep Learning

Automation required tailored structure



Data Representation

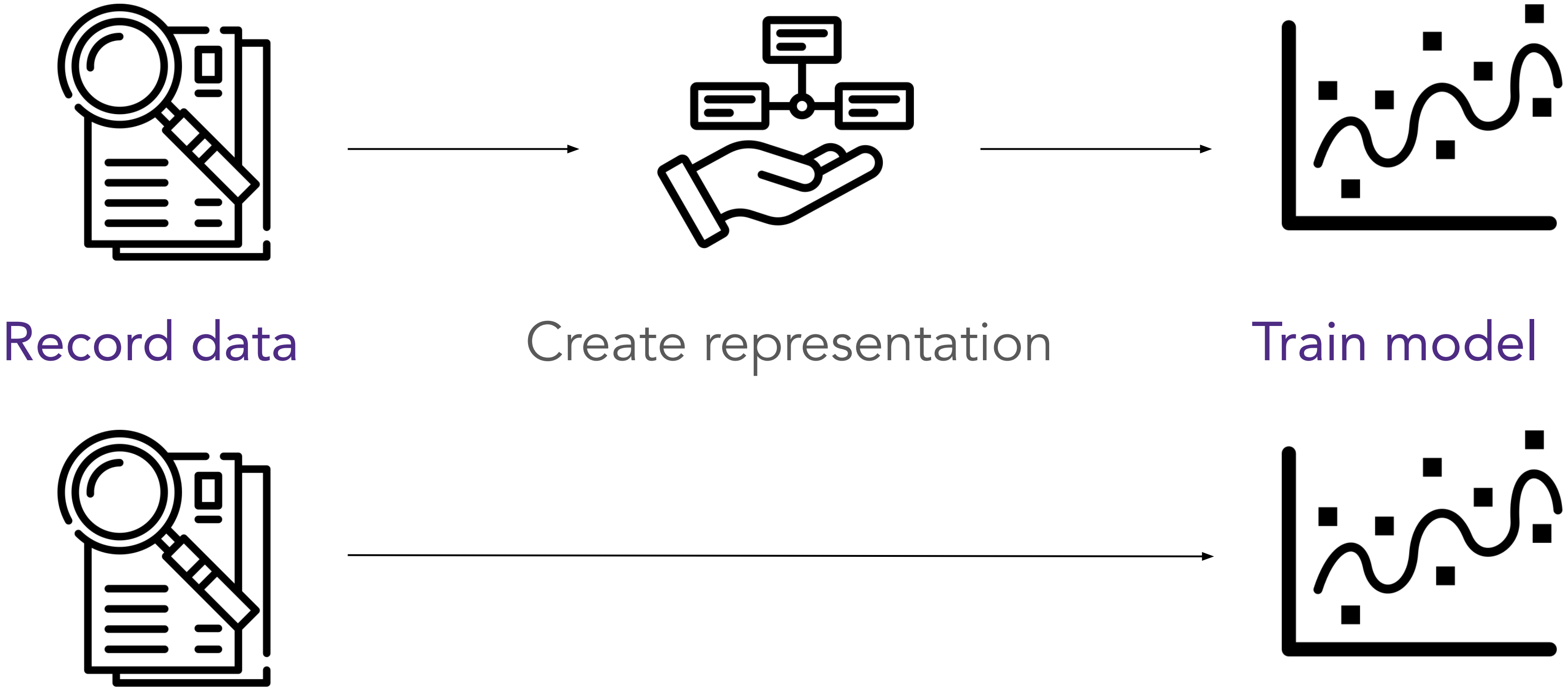
- Simple
- Structured
- Complete
-

Task Knowledge

- Verification
- Simulation
- Specification

Chess

Deep learning does the feature engineering



A machine making simple observations

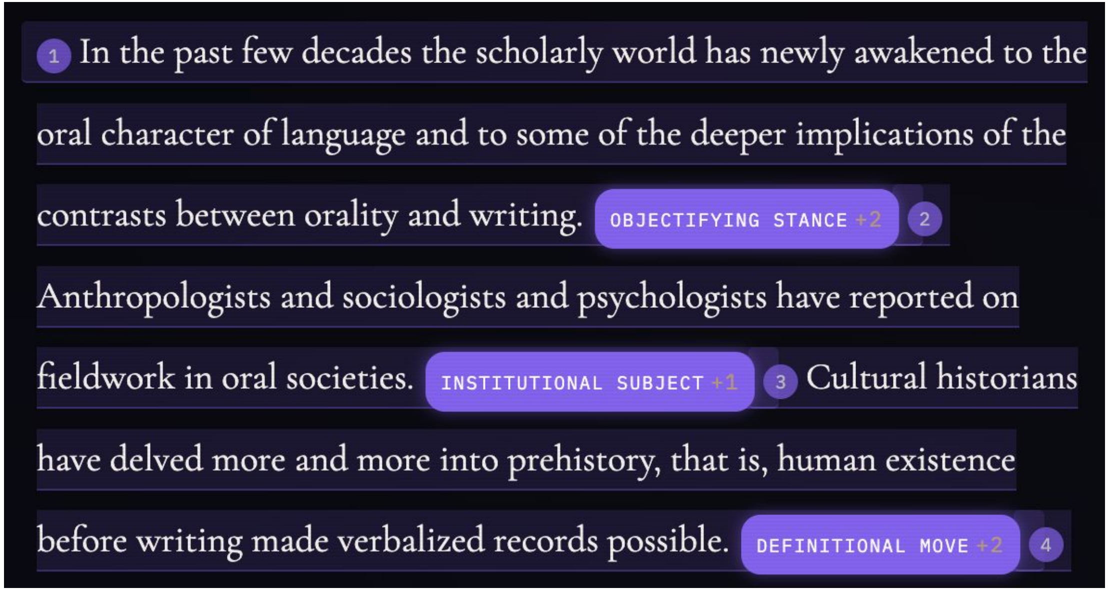
An expanded suite of capabilities

Label

Extract

Score

What is this?



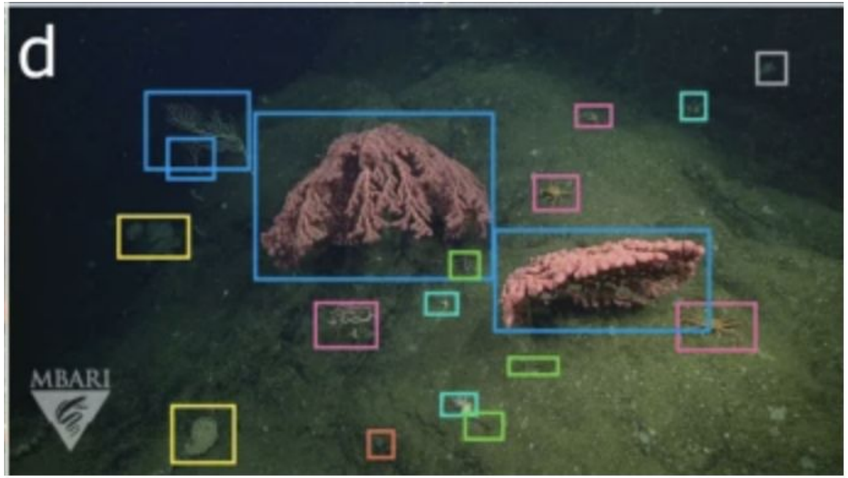
Label

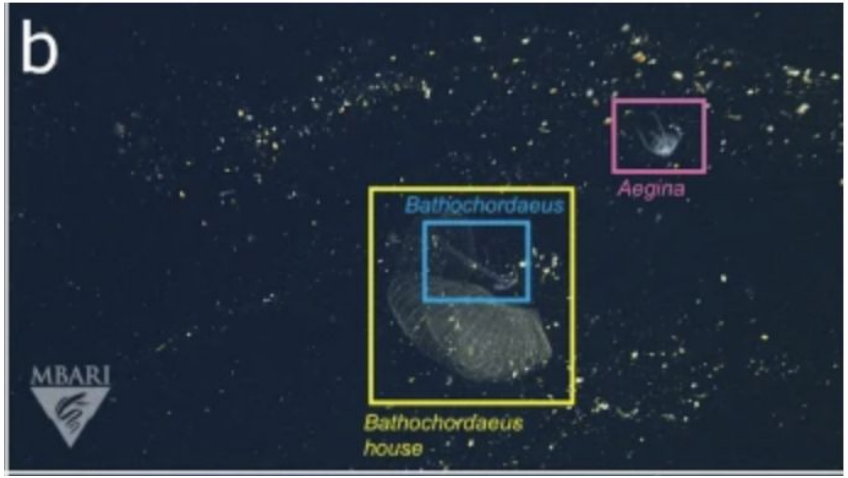
Extract

Score

From havelock.ai

Where is that thing?





Label

Extract

Score

From https://www.nature.com/articles/s41598-022-19939-2

An expanded suite of capabilities

Label

Extract

Score

From https://help.useloops.com/en/articles/7192111-what-is-the-sentiment-analysis-score

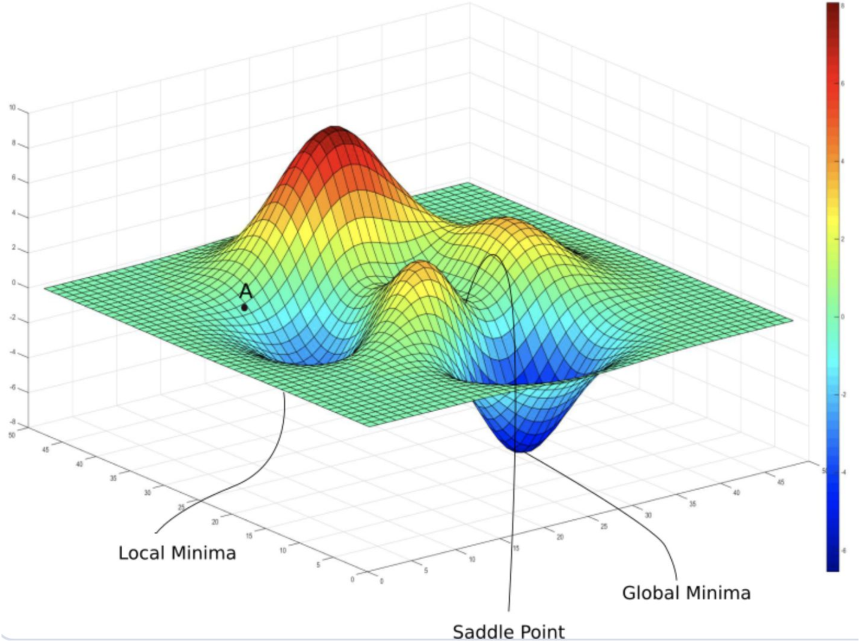
The models learn their input representations

(a) Husky classified as wolf (b) Explanation

Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.

	Before	After
Trusted the bad model	10 out of 27	3 out of 27
Snow as a potential feature	12 out of 27	25 out of 27

Table 2: "Husky vs Wolf" experiment results.



From https://arxiv.org/pdf/1602.04938 | From https://www.digitalocean.com/community/tutorials/intro-to-optimization-in-deep-learning-gradient-descent

Is deep learning the tool for that job?

Deep Learning

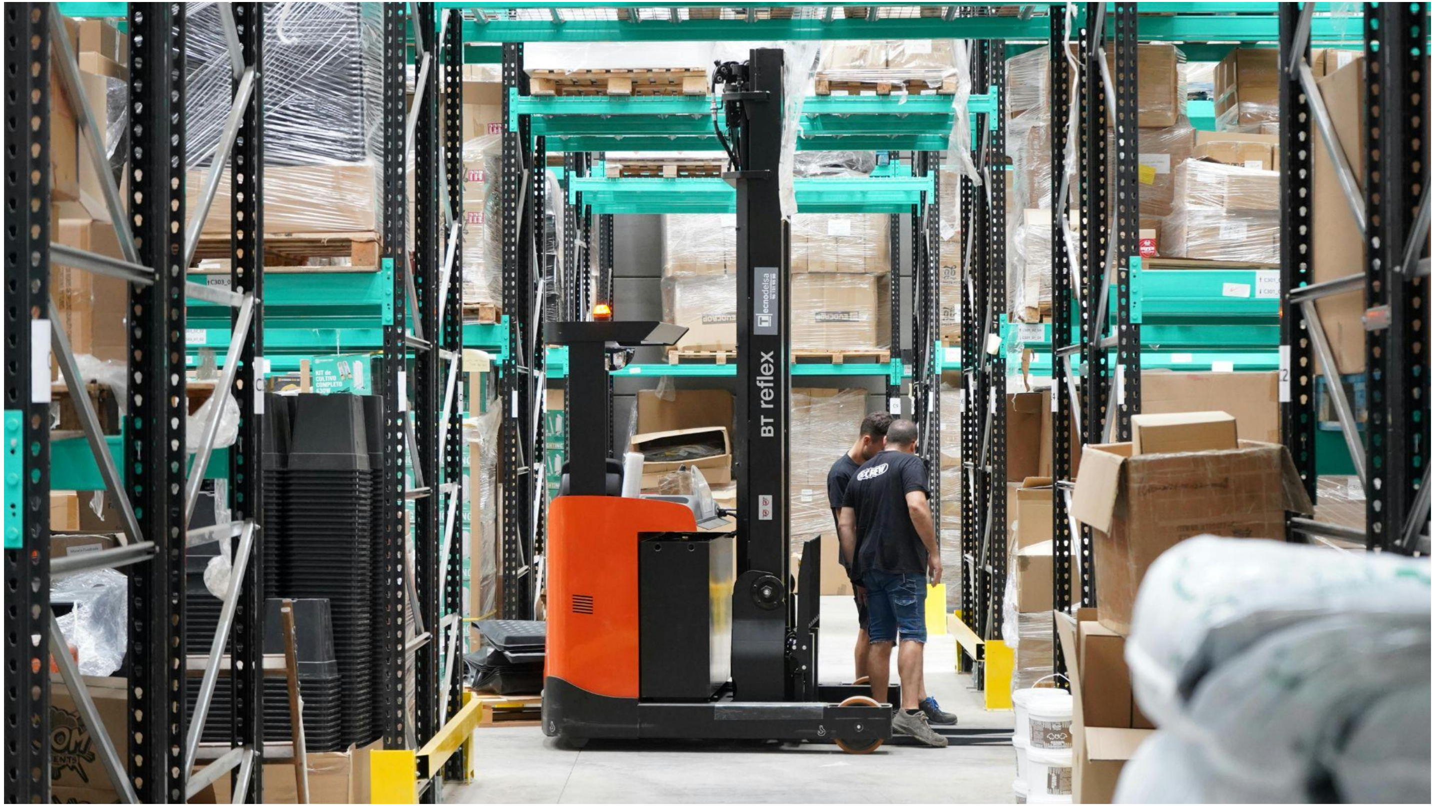
Bank branch

Grocery store

Doctor’s office

Construction site

Logistics hub



Natural preservation

Deep learning case studies

Deep Learning

Case Study: Content moderation

Problem

- Publishing platforms need to remove undesirable content, at scale

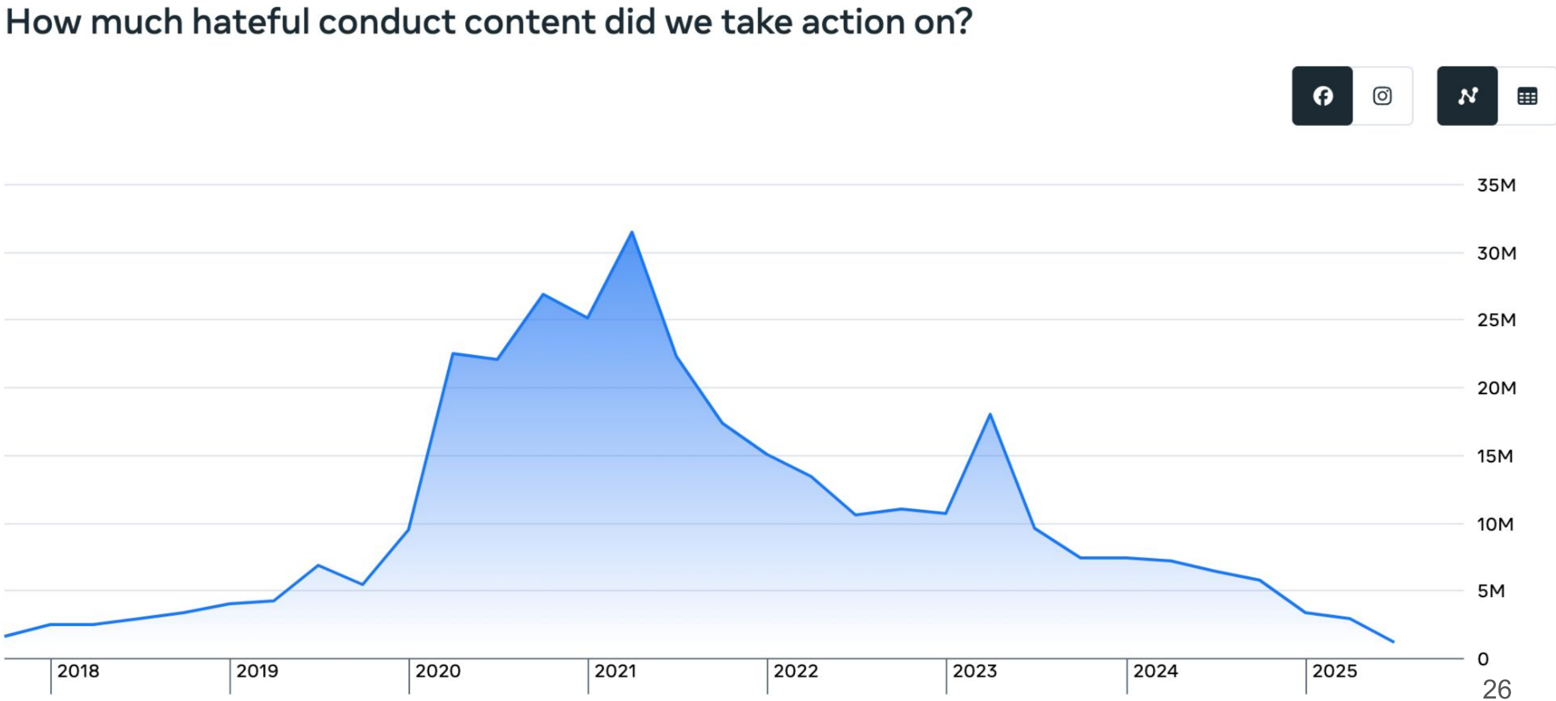
Where deep learning helps

-

Case Study: Content moderation

Where deep learning helps

-

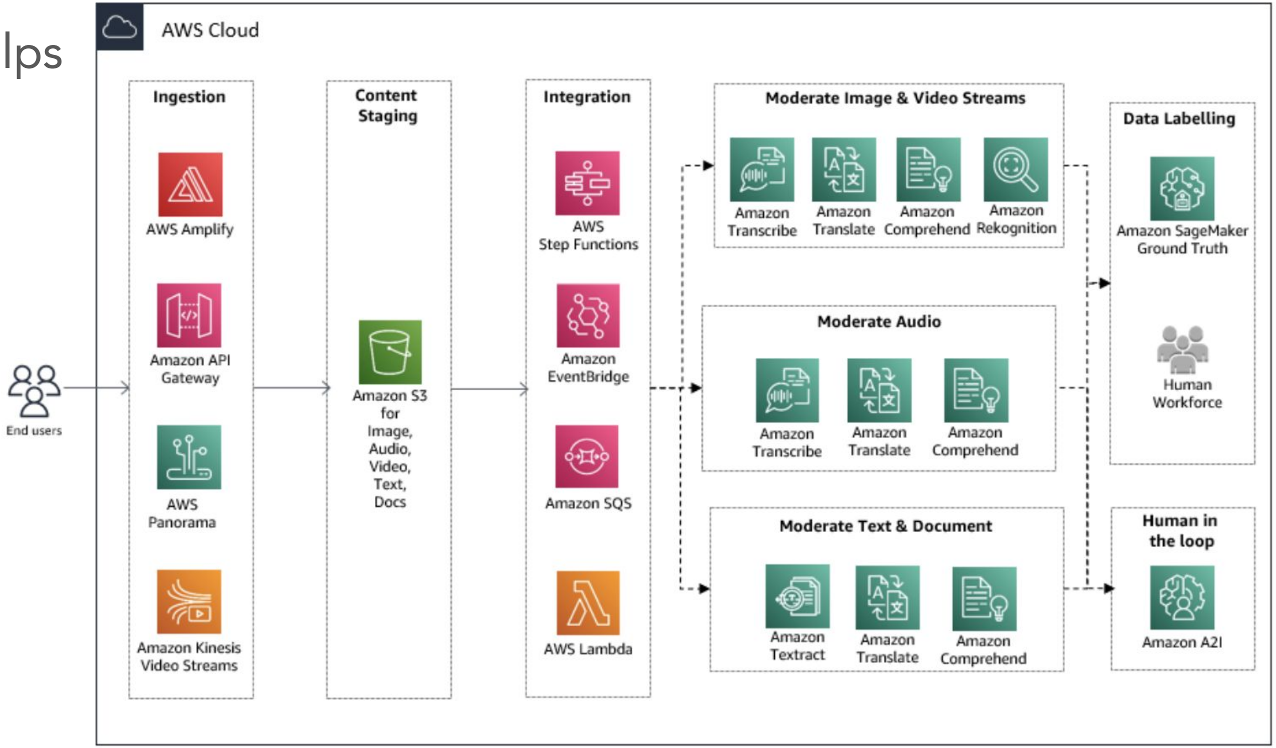


From https://transparency.meta.com/reports/community-standards-enforcement/hateful-conduct/facebook/

Case Study: Content moderation

Where deep learning helps

-



From https://aws.amazon.com/blogs/machine-learning/content-moderation-design-patterns-with-aws-managed-ai-services/

Case Study: Content moderation

Impact

- Human reviewers see fewer cases, handling exceptions
- Content can be removed immediately, but decisions may be opaque
- User behavior changes in response to moderation
- New content moderation approaches can be deployed rapidly

Case Study: Pharmacovigilance

Problem

- Analyzing the impact of medical products is complex and labor intensive

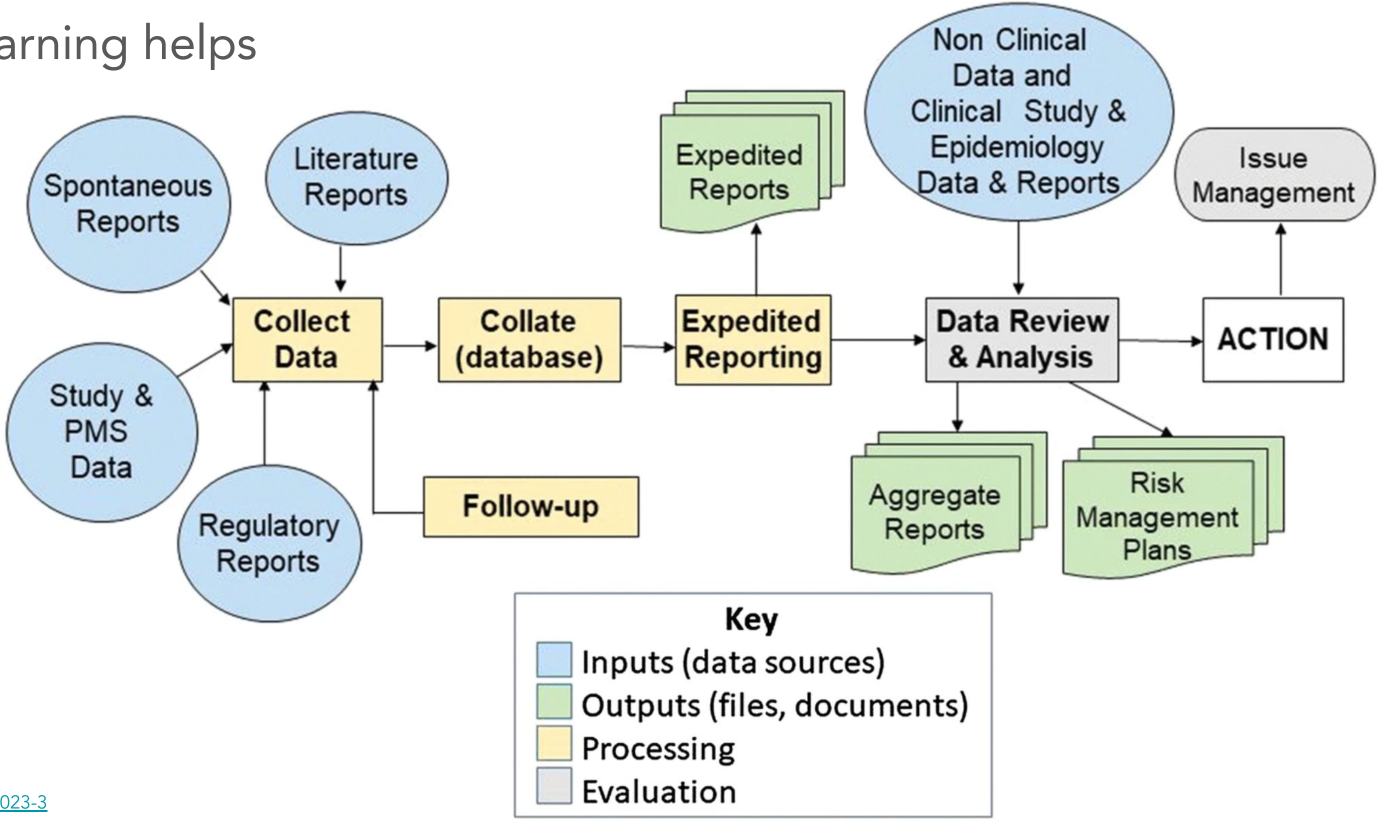
Where deep learning helps

-

Case Study: Pharmacovigilance

Where deep learning helps

-



From https://doi.org/10.1007/s43441-019-00023-3

Case Study: Pharmacovigilance

Impact

- Intake capacity expanded: can now process literature, social media at scale
- Expert human attention can be deployed more sparingly in the process
- Time-to-detection decreased for emerging signals, but more signals now
- Regulatory accountability unchanged: human sign-off by design

Case Study: Vehicle rental

Problem

- Equipment rental companies need to validate customer rentals

Where deep learning helps

-

Case Study: Vehicle rental

Where deep learning helps

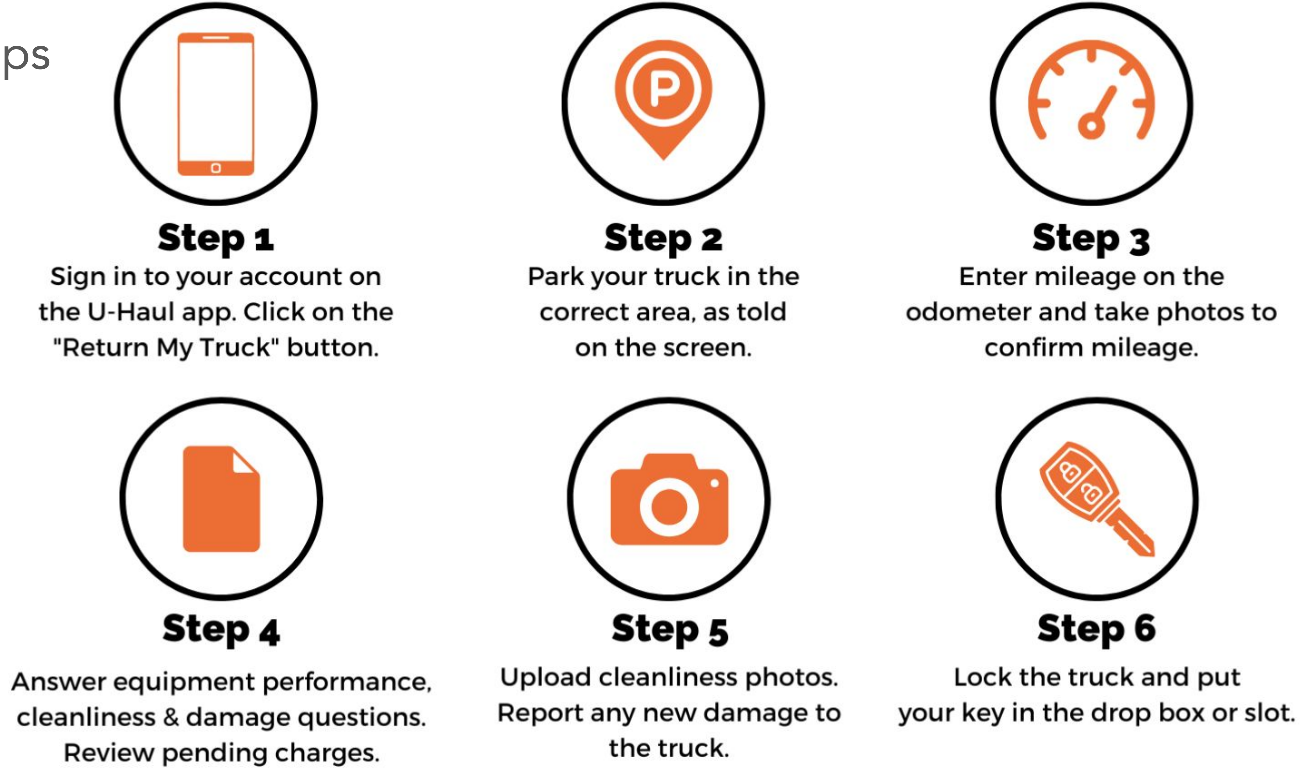
-

From https://www.facebook.com/photo.php?fbid=10157283224611094&id=91754831093&set=a.393746286093

Case Study: Vehicle rental

Where deep learning helps

-



From https://www.uhaul.com/Blog/2023/03/31/u-haul-equipment-after-hours-drop-off-made-easy/

Case Study: Vehicle rental

Impact

- Staffing model transformed: lot attendants reduced or eliminated
- Transaction speed: no waiting for inspection
- Rental app and processing capability become critical infrastructure
- Impact on site economics: do customers buy add-ons in the same way?

Ask an AI model

Deep Learning

Getting AI to think with you, not for you

Stay mindful of bias

Question solutions

Expand your perspectives

Spot weaknesses or vulnerabilities

Start with your own ideas

https://www.mccormick.northwestern.edu/magazine/fall-2025/getting-ai-to-think-with-you-not-for-you/

Assignment walkthrough

Deep Learning