The Machine Learning Taxonomy

Organizing 40+ Tasks by their Mathematical Roots

Nipun Batra

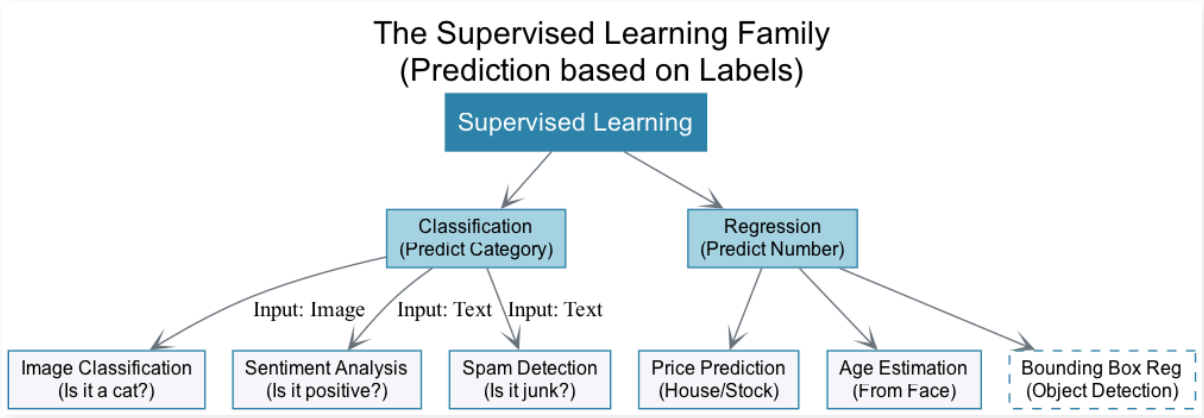
IIT Gandhinagar

The Grand Map

Machine Learning isn't just a list of random tricks. It's a family tree.
Almost every task is a variation of **Classification** or **Regression**.

- **Predicting a Label?** Classification.
- **Predicting a Number?** Regression.
- **Predicting a Sequence?** Repeated Classification.
- **Predicting a Structure?** Classification + Regression combined.

Part 1: The Supervised Family



Branch A: "Is it X or Y?" (Classification)

The simplest form of ML. $f(x) \to \{0, 1, \ldots, K\}$

1. Image Classification

`Input: Pixels | Output: Label`

Is this image a Cat or Dog?

2. Sentiment Analysis

`Input: Text | Output: Label`

Is this review Positive or Negative?

3. Spam Detection

`Input: Email | Output: Binary`

Is this junk?

4. Topic Classification

`Input: Document | Output: Category`

Is this news about Sports, Politics, or Tech?

Branch B: "How Much?" (Regression)

Predicting continuous values. $f(x) \rightarrow \mathbb{R}$

5. House Price Prediction

`Input: Features | Output: $ Price`

Classic regression.

6. Age Estimation

`Input: Face Image | Output: Years`

Predicting 25.4 years vs 25 years.

7. Time Series Forecasting

`Input: History | Output: Future Value`

Predicting tomorrow's temperature.

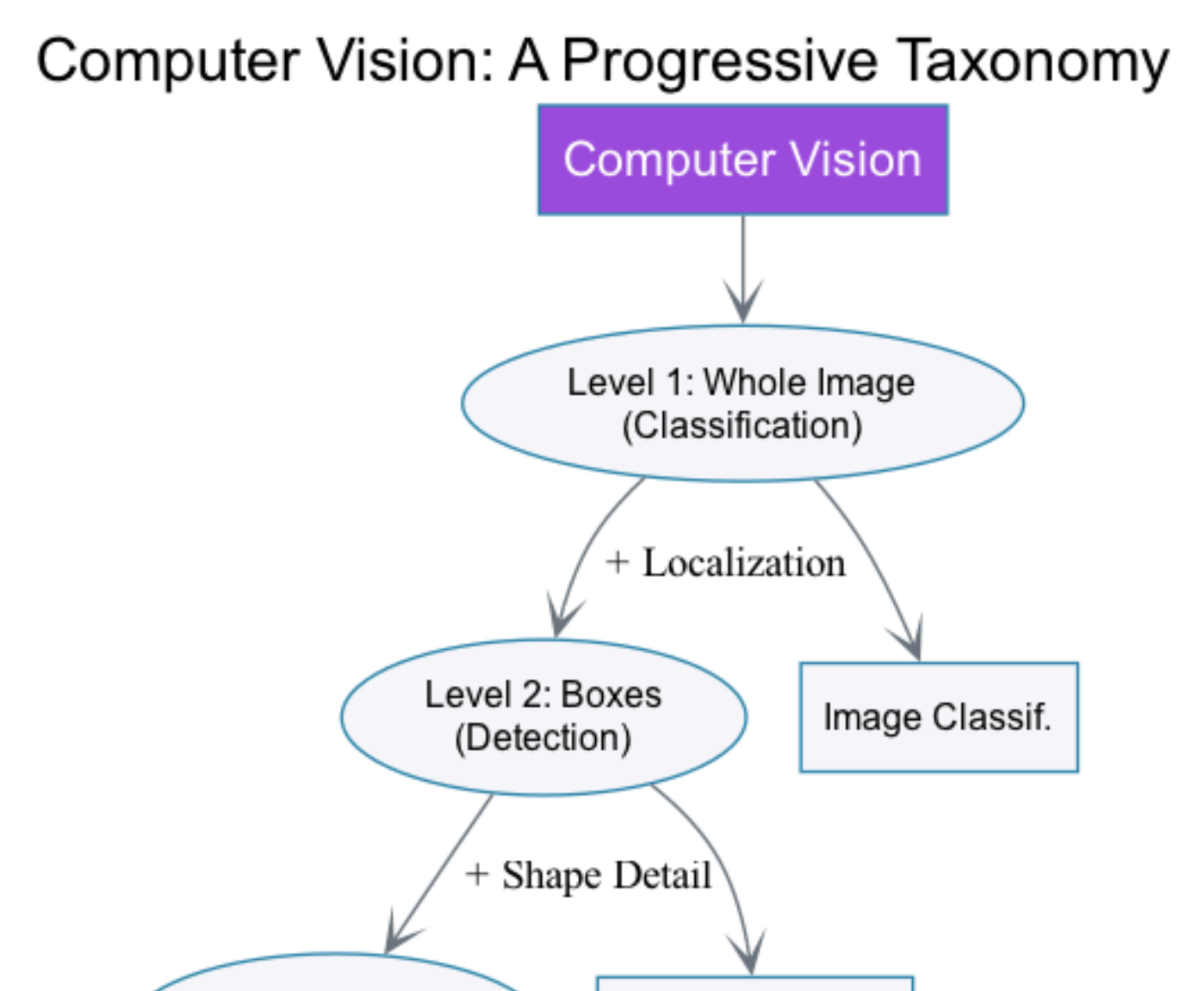
8. Bounding Box Regression

`Input: Image | Output: (x, y, w, h)`

Predicting the coordinates of an object (Part of Detection).

Part 2: The Vision Hierarchy

Combining Classification + Regression



Level 2: "Where is it?" (Detection)

We combine **Classification** (What) + **Regression** (Where).

9. Object Detection

`Output: Box + Class`

"There is a Car at [10, 50, 200, 300]"

10. Face Detection

`Output: Box`

Finding faces for auto-focus.

11. Keypoint Detection (Pose)

`Output: (x,y) points`

Finding Elbows, Knees, Eyes. (Regression of 17 points).

12. Text Detection (OCR)

`Output: Box around text`

Finding words in street signs.

Level 3: "Which Pixels?" (Segmentation)

Now we classify **every single pixel**.

13. Semantic Segmentation

`Output: Class per pixel`

Road, Sky, Tree (No distinction between two trees).

14. Instance Segmentation

`Output: Class + ID per pixel`

Car #1 vs Car #2.

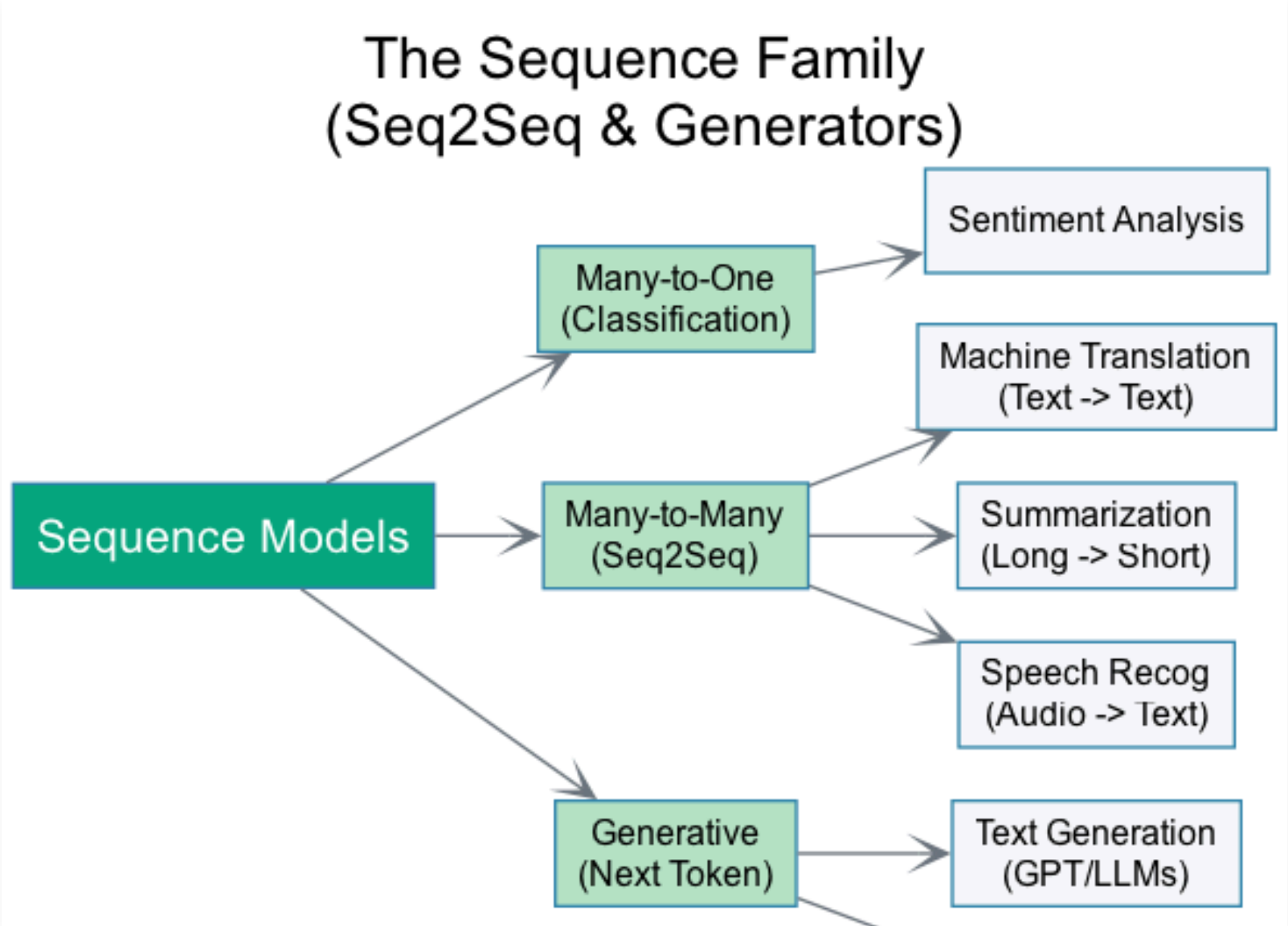
15. Image Matting

`Output: Alpha Matte (Transparency)`

Zoom background blur / Green screen removal.

Part 3: The Sequence Family

Predicting Lists of things



Many-to-Many (Seq2Seq)

Standard Classification is Many-to-One.
Seq2Seq is **Text In, Text Out**.

16. Machine Translation

`English $\rightarrow$ Hindi`

Mapping sequence to sequence.

17. Text Summarization

`Long Text $\rightarrow$ Short Text`

Extracting key information.

18. Speech Recognition (ASR)

`Audio Wave $\rightarrow$ Text`

Mapping sound frames to phonemes/words.

19. Text-to-Speech (TTS)

`Text $\rightarrow$ Audio Wave`

The reverse of ASR.

Token-Level Tasks (Tagging)

Classifying each token in a sequence (like Semantic Seg for text).

20. Named Entity Recognition (NER)

`Output: [PER, LOC, ORG] per word`

Identifying "Sundar Pichai" as PER.

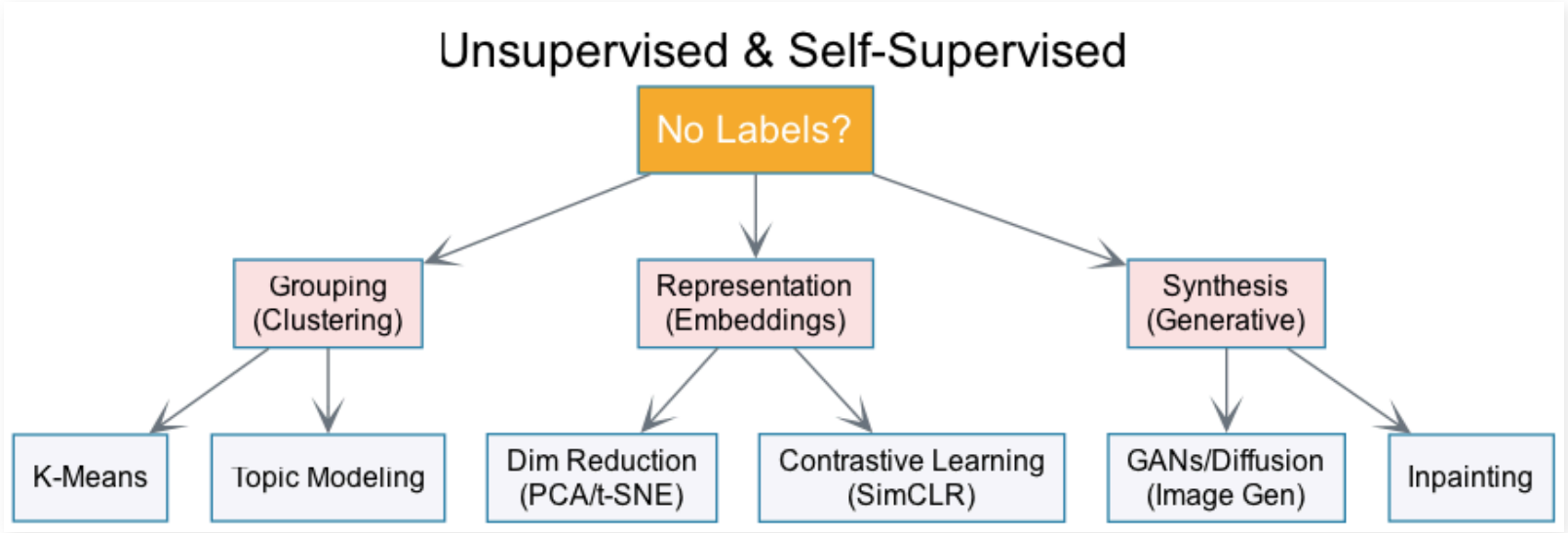
21. Part-of-Speech Tagging

`Output: [Noun, Verb] per word`

Grammatical analysis.

Part 4: Unsupervised & Generative

Learning without Labels



Grouping & Representation

22. Clustering

`Task: Find Groups`

Customer segmentation.

23. Topic Modeling

`Task: Find Themes`

Discovering "Sports" cluster in news without labels.

24. Dimensionality Reduction

`Task: Compression`

PCA / t-SNE. Visualizing high-dim data.

25. Anomaly Detection

`Task: Find Outliers`

Credit fraud, manufacturing defects.

Generative Tasks (The New Wave)

Modeling the distribution $P(X)$.

26. Image Generation

`Noise $\rightarrow$ Image`

GANs, Diffusion (Midjourney).

27. Text Generation

`Prefix $\rightarrow$ Continuation`

LLMs (GPT).

28. Inpainting

`Masked Image $\rightarrow$ Full Image`

Filling holes.

29. Style Transfer

`Content + Style $\rightarrow$ Image`

Artistic filters.

Part 5: The Complex Ones (Multimodal + RL)

30. Visual QA (VQA)

`Image + Text $\rightarrow$ Text`

"What color is the car?"

31. Image Captioning

`Image $\rightarrow$ Text`

"A dog running on grass."

32. Reinforcement Learning

`State $\rightarrow$ Action`

Playing Chess, Robot Control.

33. Recommendation

`User History $\rightarrow$ Item Rank`

Netflix/Amazon.

Summary

We didn't just list tasks. We grouped them by their **Mathematical Nature**.

1. **Classification:** The parent of Vision/NLP classification.
2. **Regression:** The parent of Prediction/Bounding Boxes.
3. **Seq2Seq:** The parent of Translation/Speech.
4. **Generative:** The parent of GPT/DALL-E.

Understanding the root helps you solve the leaf.

Questions?