An Introduction to Prototype Networks

Fakir Aaqil Parvez
22BAI1092

What is Few-Shot Learning?

- Definition:
  - Few-shot learning aims to train models to generalize to new tasks with only a few labeled examples.
- Challenges:
  - Scarcity of labeled data.
  - Requires models to generalize quickly to unseen classes.
- Examples:
  - Recognizing a new object after seeing 1–5 examples.
  - Personalizing a voice assistant with minimal data.

Brief Introduction to Meta-Learning

- Definition:
  - Meta-learning, or "learning to learn," trains models to learn efficiently from limited data.
- How it works:
  - Models are trained over a distribution of tasks, not a single task.
- Goal:
  - Generalize to unseen tasks by leveraging task distribution knowledge.
- Analogy:
  - Humans learning a new skill by relying on prior experience with related skills.

Few-Shot Learning and Meta-Learning

- Key Components:
  - Task Distribution: Multiple tasks sampled for training.
  - Support Set: Small labeled dataset for learning.
  - Query Set: Dataset for evaluation on the task.
- Objective: Quickly adapt to new tasks using minimal data from the support set.

Prototypical Networks

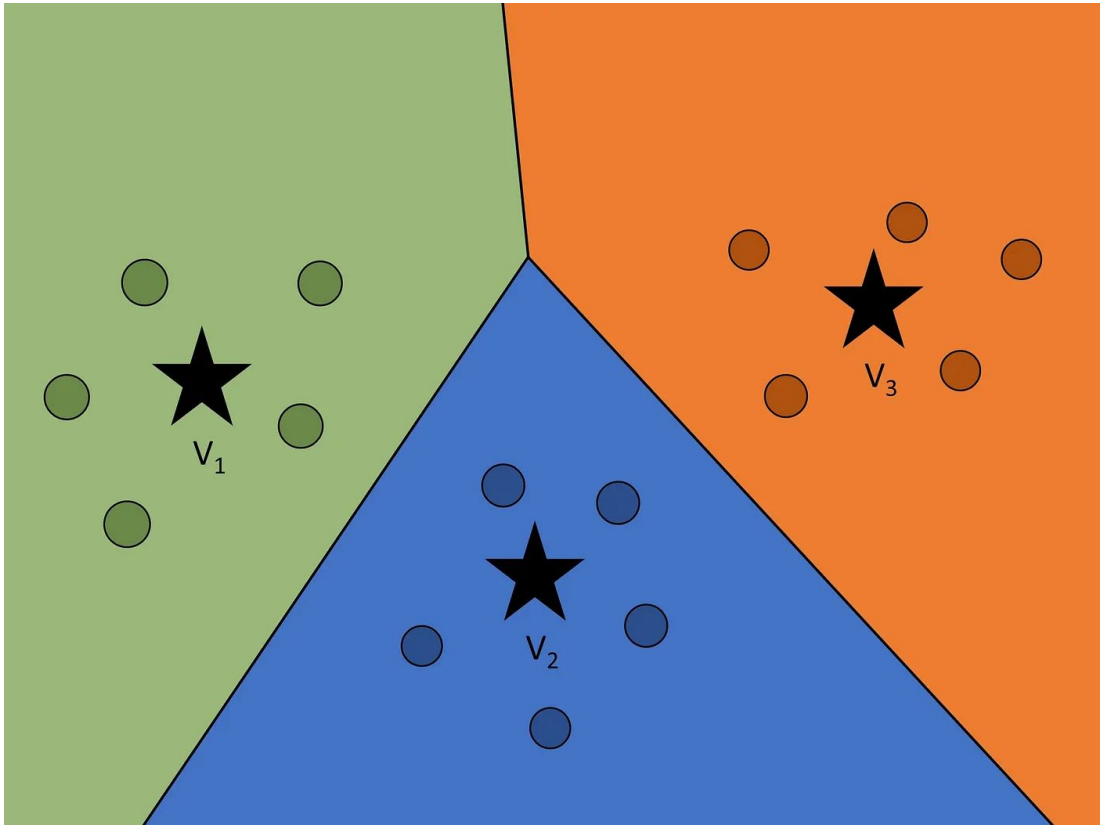
- Idea: Represent each class by a single prototype (mean of embeddings for the class).
- Why it works:
  - Assumes data points form clusters in embedding space.
  - Classification is based on proximity to these prototypes.
- Key Advantage: Simple yet effective for few-shot learning.

How Prototypical Networks Work

- Embedding Space:
  - Learn a feature extractor (e.g., CNN) to map inputs to an embedding space.
- Prototypes:
  - Compute the mean of embeddings for each class in the support set.
- Distance-Based Classification:
  - Classify query points by measuring distance (e.g., Euclidean) to prototypes.
- Loss Function:
  - Cross-entropy loss on distances ensures closer embeddings for the correct class.

Prototypical Networks in Action

- Training:
  - Tasks sampled with n_way classes and k_shot examples per class.
  - Optimize embeddings by minimizing the classification loss for query samples.
- Inference:
  - Compute prototypes for unseen tasks.
  - Classify query points based on nearest prototype.


V_1
V_2
V_3

Advantages of Prototypical Networks

- Simplicity: No need for complex memory modules or meta-optimizers.
- Efficiency: Only requires computing means and distances.
- Few-Shot Generalization: Performs well with minimal data.

Limitations of Prototypical Networks

- Assumption of Clustered Data:
  - Assumes data points of the same class form compact clusters.
- Fixed Metric:
  - Relies on a predefined distance metric (e.g., Euclidean).
  - Performance might degrade if the assumption fails.

Applications of Few-Shot Learning

- Real-World Use Cases:
  - Personalization in AI assistants.
  - Medical imaging (diagnosis with few labeled samples).
  - Species classification in biology.
  - Low-resource NLP tasks.
- Future Directions:
  - Combining with attention mechanisms, self-supervised learning, etc.