What is few-shot learning?

Few-shot learning is a machine learning framework in which an AI Model learns to make accurate predictions by training on a very small number of labelled examples. It’s typically used to train models for classification tasks when suitable training data is scarce.

Few-shot learning (FSL) is a subset of what is sometimes referred to more generally as *n-shot learning*, a category of artificial intelligence that also includes *one-shot learning* (in which there is only one labelled example of each class to be learned) and *zero-shot learning* (in which there are no labelled examples at all). While one-shot learning is essentially just a challenging variant of FSL, zero-shot learning is a distinct learning problem that necessitates its own unique methodologies.

While powerful, supervised learning is impractical in some real-world settings: obtaining labelled examples is often difficult due to prohibitive costs, the domain-specific expertise needed to annotate data correctly

How does few-shot classification work?

Though few-shot learning can utilize a wide variety of algorithms or neural network architectures, most methods are built around *transfer learning* or *meta learning* (or a combination of both).

Transfer learning-based methods focus on adapting a pre-trained model to learn new tasks or previously unseen classes of data.

When few labelled samples are available, using supervised learning to train a model from scratch—especially one with a large number of parameters, like the CNN typically used in computer-vision or the transformer-based networks used in natural language processing (NLP)—often leads to overfitting. The model might perform well on test data, but poorly on real-world data. But gathering a sufficiently large amount of data to avoid overfitting is often a bottleneck in model training.

Transfer learning offers a practical solution: leverage useful features and representations that a trained model has already learned. One simple approach is to fine-tune a classification model to perform the same task for a new class through supervised learning on a small number of labelled examples. More intricate approaches teach new skills through the design of relevant downstream tasks–often meta learning tasks—to a model that been pre-trained via self-supervised pretext tasks. This is increasingly common in NLP, particularly in the context of foundation models

Literature Survey

1. Few Shots Learning for E2E Automatic Speech Recognition

* Dhanya Eladath
* Few-Shot Learning (FSL) offers several advantages
* a paradigm shift from data-hungry deep learning methods,
* enabling learning from a few examples per class during inference similar to human learning.
* FSL methods, which belong to the class of meta-learning frameworks, leverage prior knowledge acquired from similar tasks to quickly learn a new task using very few shots per class.
* This framework is particularly advantageous in cross-domain training-inference scenarios, where efficient transferable models or embedding functions learned from a large training corpus in one domain can be used as prior knowledge to perform few-shot inference in a different domain, potentially without any fine-tuning on target domain data.

1. A survey on few-shot learning in Natural Language Processing

* Mengde Yang
* Few shots learning can be of three types Mode Based, Metric based and optimization based.
* Transfer learning: apply knowledge or patterns learnt in a different task to different but related fields or problems.
* Transfer annotated data or knowledge from relevant fields or problems.
* Variations of TL in FSL – MLTL - > effectively learn from a few target samples under reasonable assumptions.
* Limitations of Transfer learning - > Negative migration, Under adaptation, Underfitting, overfitting,
* Metal learning is used so as to realize fast learning and how fast learning and gradient descent should be more accurate.
* KNOWLEDGE DISTILLATION – soft targets associated with Teacher networks. Encompasses of Logits and characteristic distillation method. Enhances modern recommendation systems

**Few shots learning in NLP**

* Distant supervision – If two entities have a relationship in knowledge base, all sentences with these two entities (words for instance), will express relationship in some way.
* Distant supervision way to build dataset - > web encyclopaedia – network is the knowledge base – delete relationship If < threshold – annotator marks -ve if sentence is incomplete
* Learning from Distance supervision dataset - >
* Attention generator > generates class-based attention size by combining source pool distributed features and support sets and then employ ridge regression to correct deviation of word importance.
* Ridge tractor > Constructs lexical representation of the values, then predicted on query set

Challenges of Few shots learning in NLP –

* Realization of Language intelligence is too difficult.
* Cannot achieve language capability of children aged 3 or 4

**FINAL AIM: 1) To Use Few shots learning to perform Sentiment Analysis on Indic Languages such as Kannada, Telugu, Tamil, Hindi, Bengali etc.**

1. **To extend the model to translation and transliteration also. (Keyword pairing)**
2. **To also apply One-Shot Learning and Zero-shot Learning on the same models.**
3. **Training, Testing, Validation with different datasets (through some Kaggle dataset or IITM challenge datasets)**
4. **To try fitting the tasks above for languages and dialects with scantily available datasets such as Sankethi/ Havyaka Kannada**