**Alisha Iqabl**

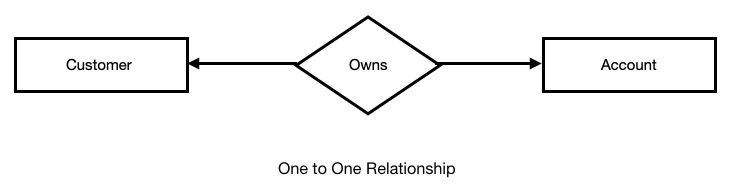
**Q1. Define the following sequence model types with diagrams:**

**i. One-to-One**

**Definition**: A traditional feed forward neural network where one input produces one output.

**Example Use Case**: Image classification   
(e.g., Classifying a single image as cat/dog)

**Real-World Example:** Handwritten digit classification (MNIST)

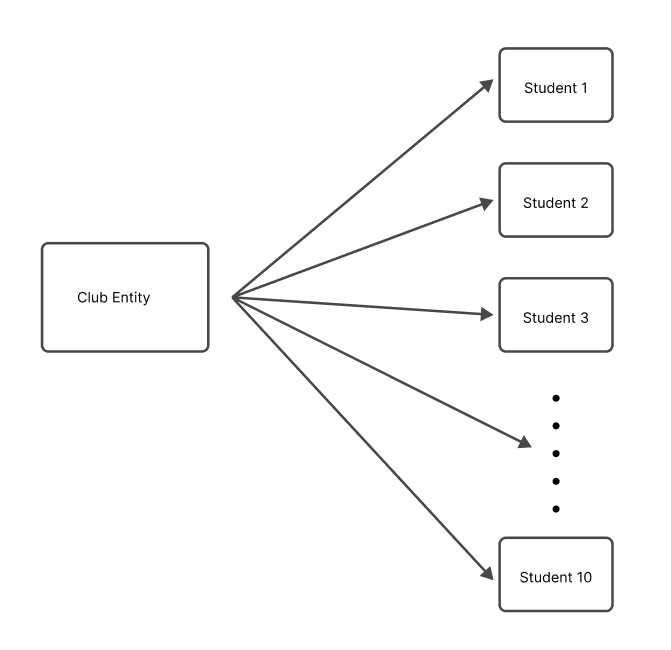


**ii. One-to-Many**

**Definition**: A model where one input leads to a sequence of outputs.

**Example Use Case**: Image captioning   
(e.g., Input: Image → Output: "A man is riding a horse.")

**Real-World Example:** Generating music from a theme



**iii. Many-to-One**

**Definition**: A sequence of inputs produces a single output.

**Example Use Case**: Sentiment analysis   
(e.g., Input: A sentence → Output: Positive/Negative sentiment)

**Real-World Example:** Spam detection in email text

#### Introduction of ER Model | GeeksforGeeks

#### ****iv. Many-to-Many****

**Definition**: A sequence of inputs produces a sequence of outputs.

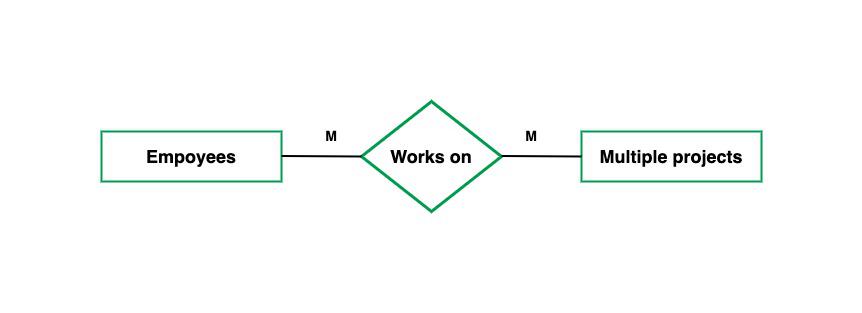
There are two types:

* **Synchronous** (input and output lengths match)
* **Asynchronous** (input and output lengths differ, e.g., in translation)

**Example Use Cases**:

* **Synchronous**: Part-of-speech tagging
* **Asynchronous**: Machine translation  
  (e.g., Input: "Bonjour" → Output: "Hello")

**Real-World Example:** Real-time speech-to-text transcription



**Q2: Scenario-Based Classification Classify each of the following tasks into one of the four types of sequence models. Justify your answer.**

**1.** Predicting the emotion of a spoken sentence.

**2.** Translating a sentence from English to French.

**3.** Predicting the next 5 words based on a keyword.

**4.** Classifying a review as positive or negative.

**5.** Tagging each word in a sentence with its part-of-speech (POS)

| **Task** | **Model Type** | **Justification** |
| --- | --- | --- |
| **1.** Predicting the emotion of a spoken sentence | Many-to-One | The input is a sequence (spoken sentence, possibly audio or text tokens), and the output is a **single label** representing the emotion (e.g., happy, sad, angry). |
| **2.** Translating a sentence from English to French | Many-to-Many (Asynchronous) | The input is a sequence (English words), and the output is another sequence (French words), often of different length. An **encoder-decoder** structure is typically used. |
| **3.** Predicting the next 5 words based on a keyword | One-to-Many | A **single keyword** (one input) generates a sequence of outputs (5 words), making it one input to many outputs. |
| **4.** Classifying a review as positive or negative | Many-to-One | The input is a sequence (text review), and the model outputs a **single sentiment label** — positive or negative. |
| **5.** Tagging each word in a sentence with its part-of-speech (POS) | Many-to-Many (Synchronous) | Each input word in the sentence gets a corresponding output POS tag, making it a **one-to-one mapping across the sequence** (synchronous Many-to-Many). |

**Q3**

**a. Why can’t we use a One-to-One model for language translation?**

**Answer:** A **One-to-One** model accepts a single input and produces a single output. However, **language translation** involves converting a **sequence of words (input language)** into a **sequence of words (target language)**, and the lengths of these sequences can vary.

**Reasons why One-to-One fails:**

* **Fixed length**: It cannot handle variable-length input and output sequences.
* **No temporal context**: It doesn’t retain or process the order or dependencies between words.
* **No mapping of sequences**: Translation requires understanding sentence context and generating words sequentially — which One-to-One cannot do.

**b. What are the key challenges of Many-to-Many sequence modeling?**

**Answer:**

Many-to-Many models are powerful but come with several challenges:

1. **Variable sequence lengths**
   * Input and output sequences can differ in length (e.g., translation), making alignment difficult.
2. **Long-term dependencies**
   * Understanding distant relationships in a sequence is hard (especially for RNNs/LSTMs).
3. **Memory and computation**
   * Training and inference are resource-intensive for long sequences.
4. **Alignment**
   * Matching which input tokens contribute to which output tokens can be non-trivial (handled by attention mechanisms).
5. **Data sparsity**
   * Complex sequences require large, high-quality datasets.
6. **Beam search complexity**
   * For decoding (e.g., translation), simple greedy output isn’t enough — complex methods like beam search are used, increasing complexity.

### ****c. Compare and contrast One-to-Many and Many-to-One using an example of weather forecasting****

| **Feature** | **One-to-Many** | **Many-to-One** |
| --- | --- | --- |
| **Definition** | Single input → Sequence output | Sequence input → Single output |
| **Weather Example** | Given today’s weather → Predict the next 5 days’ weather | Given past 7 days’ weather → Predict if it will rain tomorrow |
| **Use Case** | Forecasting future weather conditions (e.g., temperature trend) | Classification/Decision task (e.g., rain/no rain) |
| **Sequence Direction** | Output is time-dependent | Input is time-dependent |
| **Complexity** | Temporal output handling | Temporal input understanding |