

## Sparse

Def: A method of representing data as a vector or matrix where most elements are 0. In other words, it represents only some of the important features of the data and fills the rest with 0s.

### Characteristics:

**Memory Efficiency:** Memory usage is low because only non-zero elements are stored. This is especially effective for high-dimensional data.

**Computational Efficiency:** Computation speed is fast because operations with 0 can be omitted.

**Ease of Interpretation:** It is easy to clearly identify which features are important.

- Examples:
- \* Bag-of-Words (BoW):When representing text data by the frequency of words, words that do not appear in the document are treated as 0.
  - \* One-Hot Encoding:\*When representing categorical data, only the element corresponding to the category is marked as 1, and the rest are treated as 0.
  - \* TF-IDF (Term Frequency-Inverse Document Frequency):\*\* Weights are assigned to words considering the frequency of words in a document and the importance in the overall document, and unimportant words are represented with values close to 0.

### Disadvantage

**Information Loss:** Because many pieces of information are treated as 0, subtle features of the data may be missed.

**Limited Expressiveness:** It can be difficult to effectively represent the relationships between data.

## Dense

Def: AA method of representing data where all elements of a vector or matrix are **filled with non-zero values**. All features of the data are represented by real values.

### Characteristics:

**Rich Information:** More information can be contained by utilizing all features of the data.

**Improved Expressiveness:** Relationships between data can be effectively represented.

**Ease of Interpretation:** It is easy to clearly identify which features are important.

- Examples:
- Word Embedding:Represents words as vectors in a high-dimensional space, reflecting the semantic similarity between words (Word2Vec, GloVe, FastText, etc.)
  - Image Pixel Values:\*When representing an image with pixel values, all color information of each pixel is used.
  - Hidden Layer Output of Neural Networks:\*\* The hidden layer of a neural network transforms input data into a dense vector to learn more complex features.

### Disadvantage

- \* \*\*Memory Inefficiency:\*\* Memory usage is high because all elements must be stored.
- \* \*\*Computational Complexity:\*\* Computation speed is slow because operations are required for all elements.
- \* \*\*Difficulty of Interpretation:\*\* It can be difficult to identify which features are important.