Lab 3 : Theoretical Task

D7047E, Advanced Deep Learning

Submitted by : Tan Wei Zhong , Sandra Ljungholm , Irek Sakhabutdinov, Veronica Baranova, Harish Vasanth

Theoretical task :

Explain the pros and cons of using Concatenation , Addition , Multiplication , Attention , Difference for the process of combining embeddings.

Concatenation:

Pros:

- Rich Information Fusion: Concatenation combines both embeddings into a single vector, preserving all information from both modalities.
- Simple Implementation: It's straightforward to implement by stacking the embeddings together.

Cons:

- Fixed-size Representation: The combined vector might become too large and may not effectively capture the relationships between the two modalities.
- Potential Redundancy: Some redundant information from one modality might dilute the representation of the other.

Addition:

Pros:

- Simplicity: Addition is straightforward and computationally efficient.
- Preserves Original Scale: Retains the scale of the original embeddings, potentially balancing the contribution of each modality.

Cons:

- Loss of Discriminative Information: Addition might oversimplify the relationship between embeddings, potentially masking important information.
- Sensitivity to Scale: The addition operation can be sensitive to the scale of the input embeddings

.
Multiplication:

Pros:

- Selective Information Fusion: Multiplication can selectively highlight or suppress certain features from each modality.
- Non-linear Interaction: This method introduces non-linear interactions between embeddings.

Cons:

- Complexity: It can introduce additional complexity in modeling due to non-linear interactions.
- Loss of Information: Might attenuate the representation of less dominant features from either modality.

Attention:

Pros:

- Selective Focus: Attention mechanisms allow the model to dynamically focus on relevant parts of each modality based on context.
- Adaptive Fusion: The model can learn to emphasize important information from each modality based on the caption generation process.

Cons:

- Complexity: Attention mechanisms require additional computational resources and may increase model complexity.
- Training Challenges: It can be more challenging to train due to the need to learn the attention weights effectively.

Difference:

Pros:

- Contrastive Information: Difference can highlight complementary aspects or unique characteristics between embeddings.
- Simplicity: Like addition, it's a relatively simple operation.

Cons:

- Sensitive to Scaling: The difference operation can be sensitive to the scaling of the input embeddings.
- Potential Loss of Context: Focusing on differences alone may overlook important relationships or similarities.