**Introduction**

Meta-learning algorithms have been very successful on new fixed data split into train, validation, and test sets, but when applied to real-world applications, data can often be distributed differently. In order to address this problem, we try to improve the robustness of our models to increase their generalizability.

**Problem**

Machine learning algorithms have shown great success in answering in-domain queries, but similar to many NLP problems, it suffers from poor generalizations to a few-shot out-of-distribution context.

Meta-learning is an approach that "learns to learn" from other NLP models. It shares prefixes in effectively generating new out-of-domain data distributions. In short, it peaks at a new out-of-context position for the model to attack a problem with fewer input data. Meta-learning has been previously explored in few-shot NLP learning tasks. For instance, Gao et al. at al. trained the replica model on the GLUE benchmark, however, meta-learning in NLP has seen less success than in other fields. We explore the possibility of implementing meta-learning to out of domain question answering.

In addition, data augmentation methods have shown improvement results in improving the robustness, particularly in computer vision. Thus, we explore various augmentation operations and investigate the best combination of operations and hyperparameters to boost performance.

**Methodology and Experiments**

**Easy Data Augmentation**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch Size</td>
<td>128</td>
</tr>
<tr>
<td>Learning Rate</td>
<td>5e-5</td>
</tr>
<tr>
<td>Epochs</td>
<td>30</td>
</tr>
</tbody>
</table>

**Easy Data Augmentation**

- **Data Processing**
  - All words lengthened by 1 character
  - Word Punctuation retained

- **EDA Operations**
  - **Syntactic Replacement**: random replacement of a word with another
  - **Random Insertion**: random insertion of words
  - **Random Deletion**: random deletion of words

**Example**

<table>
<thead>
<tr>
<th>Original Text</th>
<th>EDA Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is a sentence.</td>
<td>This is another sentence.</td>
</tr>
</tbody>
</table>

**Meta Learning**

- **Learning to learn**: effect for few-shot environments
- **Feature Transfer**: good initialization of a neural network so that the model could be easily fine-tuned on few-shot datasets

**Meta Learning Setup**

- **Algorithm**
  - **Eqs. 1**: standard meta-learning

**Data EDA**

- **EDA Text**: EDA operation applied

**Table 1**: Content generated by EDA operation, example taken from the interaction interaction dataset, each example is 100 words.

**Results**

- **EDA Text**
  - **Syntactic Replacement**: random replacement of a word with another
  - **Random Insertion**: random insertion of words
  - **Random Deletion**: random deletion

**EDA Ablation Study**

<table>
<thead>
<tr>
<th>Hyperparameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch Size</td>
<td>128</td>
</tr>
<tr>
<td>Learning Rate</td>
<td>5e-5</td>
</tr>
<tr>
<td>Epochs</td>
<td>30</td>
</tr>
</tbody>
</table>

**EDA Comparison**

- **EDA Text**
  - **Syntactic Replacement**: random replacement of a word with another
  - **Random Insertion**: random insertion of words
  - **Random Deletion**: random deletion

**Table 2**: Experimental settings used by EDA methods.

**Discussion**

**Data Augmentation**

- **Goal**
  - Few-shot scenario: in a few-shot scenario, the model must be able to generalize to new data distributions.

**Results**

- **EDA Text**
  - **Syntactic Replacement**: random replacement of a word with another
  - **Random Insertion**: random insertion of words
  - **Random Deletion**: random deletion

**Table 3**: Results from meta-learning models trained with EDA Split, Separation, and EDA by Task split.

**FUTURE DIRECTIONS**

- **Data Augmentation**
  - **Goal**
    - Few-shot scenario: in a few-shot scenario, the model must be able to generalize to new data distributions.

- **Results**
  - **EDA Text**
    - **Syntactic Replacement**: random replacement of a word with another
    - **Random Insertion**: random insertion of words
    - **Random Deletion**: random deletion

- **Table 3**: Results from meta-learning models trained with EDA Split, Separation, and EDA by Task split.