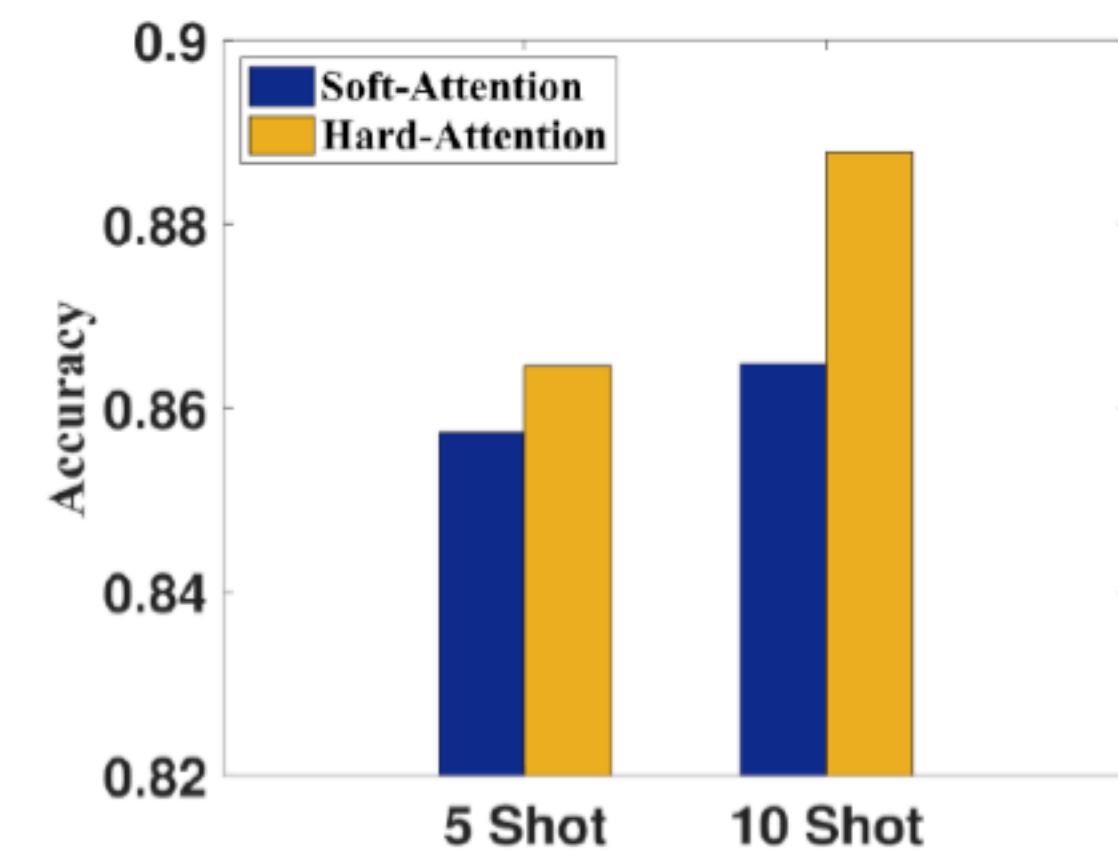


# Experiments

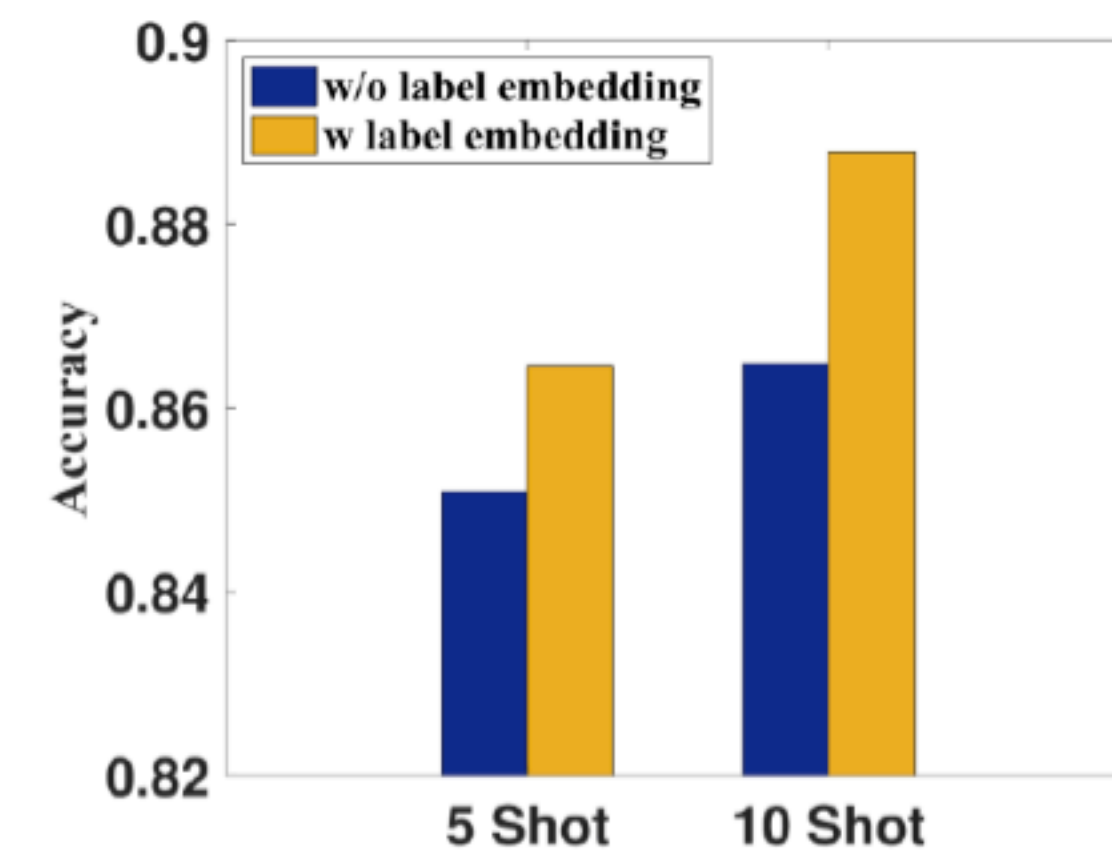
## Ablation Study: Soft-Attention v.s. Hard-Attention



- **Hard-attention** in 5 or 10-shot settings are **greater than those of Soft-attention**.
- As the number of support set increases, hard-attention mechanism doesn't have limitation of soft-attention mechanism which **unavoidably incorporates unrelated data points**.
- Conclude that hard-attention can take effectively advantage of support set, and the superiority is **more significant when enlarge size** of support set.

# Experiments

## Ablation Study: w/o v.s. w/ Label Embedding



- Reduced model by replacing label embedding with label value 0 or 1.
- Change the multiplication between output with label embedding to a binary-class fully connected layer to directly output the probabilities.
- Observe that the accuracy score of **w/ label embedding is greater** than w/o label embedding in two shot settings.
  - Demonstrating the **effectiveness** of label embedding.