Advancing Explainability/Accuracy in *ServeNet*-GPT-Fusion: Integrating Attention Mechanisms, Feature Attribution, and Model Transparency

**Introduction:**

*ServeNet* is a stacked deep neural network and sequence model for web services classification. It refers to a deep learning-based approach or model designed for the automatic classification of web services into predefined categories.

For this purpose, our previous work was targeted for improving the accuracy of the base model by *ServeNet*-GPT-Fusion, an advanced model that integrates deep learning techniques with large language models (LLMs) like ChatGPT (GPT-3.5 and GPT-4). By application of prompt engineering, the study refines web service classification through automated methods, achieving higher accuracy than the baseline *ServeNet* model. The research employs a weighted voting mechanism to combine the predictive capabilities of multiple models, thereby enhancing classification performance. The fusion architecture integrates *ServeNet* with GPT-3.5 and GPT-4. These models bring advanced natural language understanding and semantic processing capabilities. The recent model achieves higher classification accuracy than standalone models, making it ideal for handling complex, diverse datasets.

**Problems:**

**Explainability** remains a critical challenge in web service classification models, especially those built on **black-box architecture** like deep neural networks. For example, while *ServeNet* achieves high accuracy, it cannot transparently justify why a specific service is classified under a particular category.

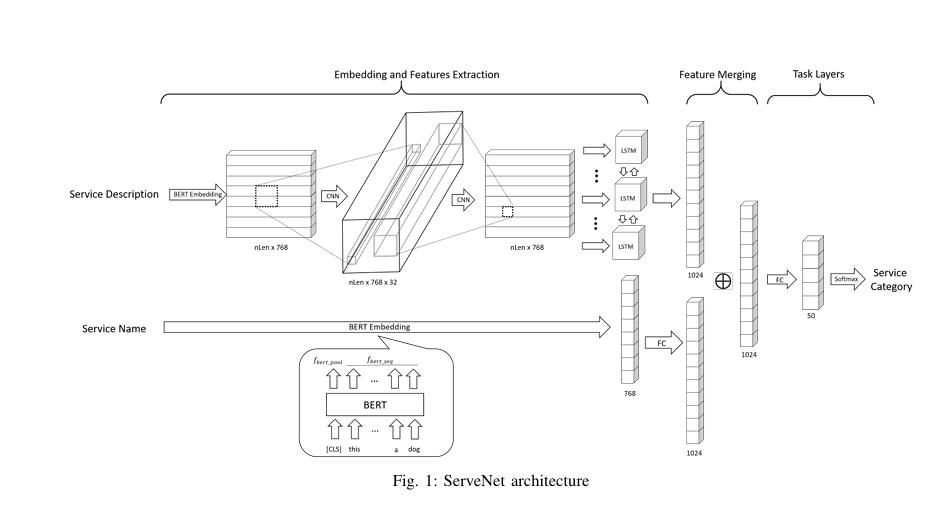
**Improving Accuracy in fusion models** *ServeNet*-GPT-Fusion

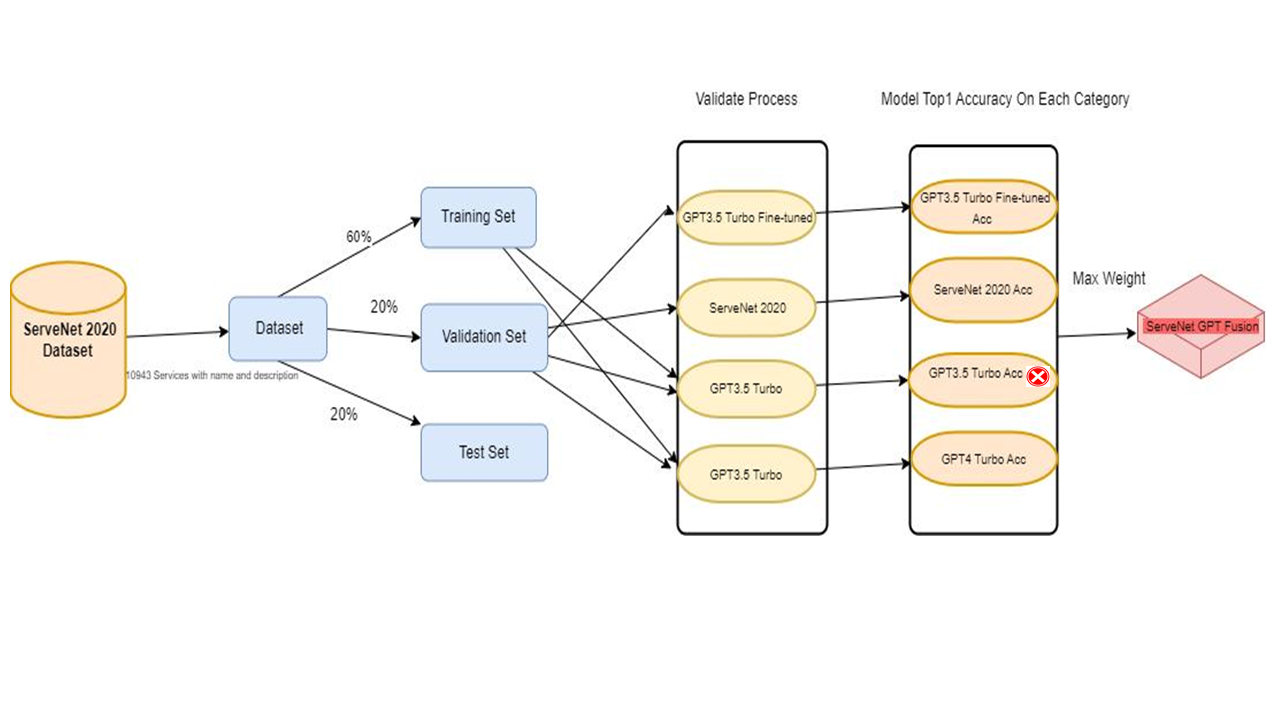
**Objectives:**

1. **Enhancing ServeNet GPT Fusion Model Accuracy with BERT + Classifier**
2. **Attention Mechanisms**: Introduce attention layers in neural networks to highlight which input features influence predictions the most. This approach makes the decision-making process more interpretable
3. **SHAP and LIME** Utilization of SHAP values or LIME to explain feature importance and the contributions of sub-models in fusion systems

**The code and original files could be found here in the folders**

**https://github.com/fred8617/ServeNet/tree/master/ServeNet-BERT https://github.com/fred8617/serveNetByGPT https://github.com/fred8617?tab=repositories**





*ServeNet*-GPT-Fusion Architecture

**Project Deliverables:**

1. **Accuracy Improvement Strategies**: - Perform hyperparameter tuning and model architecture enhancements. - Implement feature engineering techniques to improve input quality. - Consider using transfer learning or ensemble methods for better performance.

2. **New Functionalities (Optional Enhancements):** Confidence Score Implementation: Provide certainty measures for model predictions.

Explainability Features: Integrate SHAP/LIME for better interpretability. API Deployment (if required)

3. Finding top1 and top 5 accuracy on existing/new data already used in the *ServeNet*-GPT-Fusion model.

5. **Train, fine-tuning the model and handover code**

**References :**

[**https://ieeexplore.ieee.org/document/9284071/**](https://ieeexplore.ieee.org/document/9284071/) **ServeNet: A Deep Neural Network for Web Services Classification**

**https://www.semanticscholar.org/paper/ServeNet-GPT-Fusion%3A-A-Fusion-Language-Model-For-Liu-Yang/885f05c167f78f27e067efeed4b30605e3356d0b**

**ServeNet-GPT-Fusion: A Fusion Language Model For Web Services Classification**