Module 6: Critical Thinking

NLP Chatbot

Ryan Thompson

Colorado State University - Global

CSC 525

Dr. Banerjee

19 May 2025

Feature Engineering and Hyperparameter Tuning

The integration of artificial intelligence (AI) in healthcare has opened new frontiers for enhancing mental health services. Among these innovations, natural language processing (NLP) based chatbots hold particular promise for psychological therapy and diagnostic assistance. These systems simulate human-like conversation, enabling them to provide immediate, scalable support for mental health concerns. However, the effectiveness of such systems hinges on the underlying machine learning (ML) models and how well they are tuned and structured. Hyperparameter optimization and feature engineering are two critical components in this process. This paper examines, from a theoretical perspective, the machine learning models that may be used in a psychological therapy chatbot, the hyperparameters associated with these models, strategies for optimizing those parameters, and the role of feature engineering in enhancing the model’s performance.

An NLP based chatbot for psychological support and diagnosis would likely rely on deep learning architectures, with transformer based models such as Bidirectional Encoder Representations from Transformers (BERT) or Generative Pretrained Transformer (GPT) being especially suitable. These models are adept at understanding context, managing long range dependencies in language, and generating coherent text responses. For diagnostic classification tasks, such as identifying potential symptoms of depression, anxiety, or PTSD. The chatbot might employ a fine tuned transformer model, possibly in combination with other architectures like recurrent neural networks (RNNs) or convolutional neural networks (CNNs). These supplementary models can help capture patterns in sequential text data and provide additional layers of interpretability or robustness.

The prediction or classification process generally involves tokenizing the user input, converting it into numerical embeddings, passing it through the neural architecture, and generating probability distributions over potential diagnostic categories. The accuracy and generalizability of these predictions are influenced by the selection and configuration of hyperparameters.

Key hyperparameters for transformer based models include the learning rate, batch size, number of hidden layers, number of attention heads, dropout rate, and maximum input sequence length. These variables play a significant role in how effectively the model learns from data and avoids issues such as overfitting or underfitting. Improper tuning of these parameters can result in suboptimal performance or failure to converge during training.

Historically, grid search has been used to explore hyperparameter combinations. However, this method is computationally expensive and often ineffective, especially when only a few hyperparameters significantly influence performance. As Bergstra and Bengio (2012) argue, in high dimensional hyperparameter spaces, random search is often more effective. By randomly sampling from distributions rather than systematically evaluating fixed combinations, this method increases the chance of identifying impactful configurations quickly.

Bayesian optimization represents a more sophisticated approach. It builds a probabilistic model of the function mapping hyperparameters to validation performance and iteratively updates this model using new observations. This strategy balances exploration and exploitation, leading to more efficient tuning, particularly for computationally expensive models like deep neural networks (Snoek, Larochelle, & Adams, 2012). Other alternatives include the use of evolutionary algorithms, which mimic natural selection to explore the hyperparameter space over multiple generations, and reinforcement learning based methods that treat hyperparameter tuning as a sequential decision making problem (Young, Hazarika, Poria, & Cambria, 2018).

Feature engineering in the context of a psychological therapy chatbot involves transforming raw text input into informative attributes that support accurate predictions. Since mental health diagnosis depends heavily on subtle cues in language, contextual understanding, and longitudinal analysis, the importance of effective feature extraction cannot be overstated.

One significant area of feature engineering lies in linguistic analysis. Extracting features such as part-of-speech tags, sentiment polarity, word usage frequency, and syntactic complexity can offer insights into the user’s mental state. For example, frequent use of first person singular pronouns and negatively valenced adjectives has been linked to depression and other psychological conditions (Tausczik & Pennebaker, 2010). These linguistic cues can be used alongside the model’s learned embeddings to strengthen its diagnostic capabilities.

Temporal analysis also plays a vital role. By examining how language patterns evolve over time, the chatbot can detect trends indicative of mental health decline or improvement. A shift from neutral to increasingly negative sentiment across multiple interactions might flag the need for intervention. Similarly, conversational patterns such as response latency, coherence, or topic persistence can serve as behavioral markers of psychological states.

In addition, the use of semantic embeddings generated from pretrained language models enables the chatbot to capture deep contextual relationships within the text. These embeddings serve as high dimensional representations of meaning and are often used as inputs to classifiers. They allow the model to differentiate between similar phrases that carry different emotional or diagnostic implications based on context.

Other useful features might include metadata, such as the time of day or session frequency, which could correlate with specific behavioral patterns. However, the use of such information must be handled with strict attention to privacy and ethical considerations, particularly given the sensitive nature of mental health data.

Collaborating with domain experts such as psychologists is essential to ensure that the features engineered correspond to meaningful clinical constructs. Additionally, dimensionality reduction and feature selection techniques, such as principal component analysis (PCA), recursive feature elimination (RFE), or L1 regularization, can be employed to improve model efficiency and reduce the risk of overfitting.

The theoretical framework for developing an NLP based chatbot for psychological therapy and diagnosis underscores the importance of hyperparameter optimization and feature engineering. Effective tuning methods such as random search, Bayesian optimization, and evolutionary algorithms are well suited to managing the complexity of deep learning models. Simultaneously, carefully engineered linguistic, temporal, and semantic features can significantly improve the model’s performance and clinical utility. Although this discussion is theoretical, it highlights the critical considerations necessary for building robust, ethical, and effective AI systems in the domain of mental health care.

References

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13(2), 281–305.

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. In Advances in Neural Information Processing Systems, 25, 2951–2959.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology, 29(1), 24–54.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. IEEE Computational Intelligence Magazine, 13(3), 55–75.

Zhou, Z. H. (2015). A brief introduction to weakly supervised learning. National Science Review, 5(1), 44–53.