2D Project — Mathematical Modelling
Illuminating Engagement: Machine Learning Optimizes Lighting For Enhanced User Experience

Group 08 — SC05

Tsang Ho Mun Titus — 1005772

Thng Aik Kiat — 1007781

Nguyen Quoc Dung — 1008017

Ong Xuan — 1008044

Tan Sheng Zhe — 1008452

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Modelling Space & Systems

Q&A SESSION

1. What is the purpose of this MLP model, and what problem is it trying to solve?

The MLP model aims to predict user engagement levels based on lighting conditions such as light intensity, color temperature, and ambient noise. By uncovering these relationships, it enables strategic lighting adjustment and optimization for enhanced user experiences.

2. What are the input features used by the MLP model, and how are they related to the target variable (user engagement)?

The input features are light intensity (lux), color temperature (Kelvin), and noise level (decibels). These lighting factors are hypothesized to have a nonlinear influence on the target variable, user engagement, which measures user involvement and participation.

3. Can you explain the structure of the MLP model, including the number of layers, neurons in each layer, and the activation functions used?

The MLP has an input layer with 3 nodes (for the 3 input features), two hidden layers with 10 and 5 neurons respectively, and an output layer with 1 node for predicting user engagement. Sigmoid and ReLU activation functions are used.

4. How are the model parameters (weights and biases) initialized, and what optimization algorithm is used for training?

The model parameters are initialized randomly, and the mean squared error loss is minimized using gradient-based optimization techniques such as stochastic gradient descent during training.

5. What loss function is employed during the training process, and why was it chosen?

The mean squared error (MSE) loss function is used as it is a common choice for regression problems and penalizes large deviations between predicted and true engagement values.

6. How is the model's performance evaluated, and what metrics are used to assess its accuracy and generalization ability?

Model performance is evaluated using metrics such as MSE, R-squared, and visualizations such as 3D plots, contour plots, and partial dependence plots on the test set to assess accuracy, interpretability, and generalization.

7. What techniques or strategies are employed to prevent overfitting and improve the model's generalization capabilities?

Techniques such as early stopping, regularization (L1/L2), and dropout are employed to prevent overfitting and improve generalization by reducing model complexity and introducing controlled noise during training.

8. How does the MLP model handle potential outliers or noise in the input data, and what preprocessing steps should be taken to ensure data quality?

Data preprocessing steps include handling missing values, removing outliers, and normalizing/standardizing the input features to ensure the model is not unduly influenced by data quality issues or scale differences.

9. Can you explain the interpretability of the MLP model, and how do you assess the contribution or importance of each input feature?

The interpretability of the MLP is assessed through techniques such as feature importance analysis, which quantifies the contribution of each input feature to the model's predictions, enabling better understanding of the lighting-engagement relationships.

10. What are the strengths and limitations of using an MLP model for this specific problem, compared to other machine learning models or approaches?

The MLP's strengths lie in its ability to model complex nonlinear relationships, flexible architecture, and data-driven approach. Limitations include dependence on data quality, potential overfitting, and challenges in interpretability compared to simpler models.

11. How does the MLP model handle missing data or imbalanced datasets, and what strategies are employed to address these challenges?

To handle missing data, techniques such as imputation (mean, median, or model-based) or simply removing samples with missing values can be employed. For imbalanced datasets, oversampling minority classes or undersampling majority classes can help balance the data distribution.

12. Can you describe the process of hyperparameter tuning for the MLP model, and how do you determine the optimal values for parameters such as learning rate, number of hidden layers, and number of neurons?

Hyperparameter tuning involves systematically searching the hyperparameter space (e.g., number of hidden layers, neurons, learning rate) using techniques such as grid search, random search, or Bayesian optimization to find the configuration that optimizes model performance on a validation set.

13. What are the computational resources required for training and deploying the MLP model, and how scalable is the approach for handling larger datasets or real-time predictions?

Training large MLP models can be computationally intensive and may require hardware acceleration (GPUs/TPUs) or distributed computing frameworks. However, once trained, deploying the model for inference is generally efficient, making it scalable for real-time predictions with appropriate infrastructure.

14. How is the trained MLP model integrated into the overall system or application, and what steps are taken to ensure its efficient and reliable deployment?

The trained MLP model is integrated into the overall system by exposing it as a service (e.g., REST API) that accepts lighting condition inputs and returns predicted user engagement levels. Monitoring, versioning, and update mechanisms ensure reliable deployment and maintenance.

15. Have you considered incorporating additional contextual or environmental factors that might influence user engagement beyond just lighting conditions? If so, what are these factors, and how do you plan to integrate them into your model?

We have considered incorporating additional contextual and environmental factors such as temperature, air quality, background music/ambient sound, and user demographics that could influence user engagement beyond lighting conditions. To integrate these factors, we plan to collect data on these variables alongside existing lighting and engagement data, perform feature engineering to derive relevant input features (e.g., comfort levels, sound decibel/frequency metrics, demographic encodings), and concatenate them with the existing lighting features as additional input nodes to the MLP model. This expanded feature space allows capturing more intricate patterns, but care must be taken to avoid overfitting and increased complexity through regularization techniques and careful architecture/hyperparameter tuning.