**Beijing Air Quality Forecast Report**

**Machine Learning Technology in - Exercise 1**



## 1. Introduction

The project deals with the forecast for PM2.5 Air pollution in Beijing using sequential data. Air quality significantly affects public health and urban planning, which makes accurate predictions necessary. The approach involves using the recurrent nerve network (RNN) and long -term short -term memory (LSTM) models. These models are well suited for sequential addiction and can reduce problems such as disappearance gradients. The project involves pre -pricing data, designing models, experimenting with configurations and evaluating results when using RMSE.

## 2. Specific target

➔ Get RMSE <4000 on Kaggle Leaderboard.

- ➔ Test more deep learning architectures (RNN, LSTM, Gru, Hybrid).
- ➔ Explore hyper lovers: Learning speed, batch size, era, dropout, optimizes.
- ➔ Evaluate convenience technique and sequence length effects.
- ➔ Present predictions in Kagal Competition Format.

**3. Data survey**

Dataset observation

The dataset was provided by the Kagal competition and includes the quality of air and weather functions per hour. The target variable is the PM2.5 concentration.

Statistical summary

The dataset includes meteorological properties (eg temporary, dewp) missing values. The PM2.5 values vary widely, showing seasonal and temporary patterns.

Pre -proprous steps

The missing values were handled using projection and further filling. The facilities were generalized and standardized. The gradual data was converted to a sliding window for model input. The dataset was divided into training, verification and test sets.

Visualization analysis

Time series plots discovered seasonal and pollution spikes. Correlation warm maps emphasized the ratio of PM2.5 and meteorological facilities. Histogram showed diagonal distribution of PM2.5 values.

**4. Model Design**

Architectural observation

The project used several architecture: Simple RNN, LSTM (1-3 stacked layers), horror variants and hybrid LSTM+Gru models.

Architectural component

Input: Thimport Sequence. Layer: LSTM/Gru (32–128) with hidden units. Dropout for regularization. The dense layer for regression output.

Justification of design option

LSTM was chosen to handle long -term addiction. Raising and early stopping reduced overfit. Adam Optimizer was used with tund learning speeds. RMSE primary assessment was metric.

**5. Experiment Table**

| Exp. # | Architecture | Units/Layers | Optimizer | LR | Batch Size | Dropout | Epochs | RMSE |
|---|---|---|---|---|---|---|---|---|
| 1 | Simple RNN | 32 x 1 | Adam | 0.01 | 32 | 0.2 | 20 | 5600 |
| 2 | LSTM | 64 x 1 | Adam | 0.001 | 64 | 0.3 | 30 | 4200 |
| 15 | LSTM + GRU | 128 x 2 | RMSProp | 0.0005 | 128 | 0.5 | 50 | 2800 |

**6. Results and discussions**

Experiments showed clear trends: RMSE improved as teams and units increased. The LSTM model improved simple RNN. Overfit was seen in deep models without waiver. Gravel performed relative to LSTM -es with low parameters. The Hybrid LSTM+Gru model received the best RMSE (<3000) with melody. Graadient cutting and Adam Optimizer reduced the disappearance/explosion that disappeared. Vizuulation (eg vs. actually PM2.5 predicted) confirmed the model's ability to capture the trends.

**7. Conclusion**

The project made models for PM2.5 forecast. LSTM-based architecture achieved an RMSE under 3000. Challenges involved missing data, overfiting and handling shield volatility. Future work will detect attention mechanisms, include external functions such as traffic and holidays and tested advanced architecture as transformers.

**8. Github Repository Link**

https://github.com/armandkay/air-alality- forecasting

9. Kaggle Submission