LABOUR WARD OPTIMISATION

IE3100R SYSTEMS DESIGN PROJECT AY 25/26 | DEPARTMENT OF INDURTIAL AND SYSTEMS ENGINEERING

SDP Supervisor: Prof He Shuangchi | **KKH Supervisors:** Ms Pang Nguk Lan, Mr Bernard Wong, Ms Jeslyn Neo

Group 16 Team Members: Agustines Ian, Arvind Ramakrishnan, Wee Cheng Yang, Long Yiqi, Leong Ming You

1. COMPANY BACKGROUND

KK Women's and Children's Hospital (KKH) is Singapore's leading public tertiary hospital specialising in high-risk conditions, for women and children, offering comprehensive care backed by research and innovation.

2. PROBLEM STATEMENT

The Delivery Suite at KKH faces **operational challenges** due to:

- Unpredictable actual delivery dates
- Unscheduled patient arrivals
- Resource constraints

These uncertainties make it **difficult to forecast:**

- Patient admissions
- Bed occupancy
- Staffing needs

This project aims to develop a **data-driven predictive model** to estimate patient admission dates using antenatal data, enabling more proactive resource planning and improved operational efficiency.

3. PROPOSED SOLUTION

Python-Based Predictive Tool for Delivery Suite Admissions

A data-driven forecasting tool will be developed using Python to predict the estimated admission timing to the Delivery Suite. By leveraging historical admission records and key antenatal patient data, the model identifies patterns that link maternal health indicators to the onset of delivery.

Enables care teams to shift from **reactive** to **proactive** planning through:

- **Demand Forecasting:** Anticipate admission surges and optimise staffing levels
- **Resource Allocation:** Reduce potential congestion at Triage or incidence of bed crunch issues by load leveling of planned inductions
- **Risk Identification:** Flag patients with higher likelihood of early or delayed labour

Impacts:

- Enhanced operational readiness
- Better resource utilisation
- May improve patient and staff experience/satisfaction through smoother operations

4. METHODOLOGIES

1. Data Processing & Cleaning

Raw data from Excel was cleaned and transformed into usable inputs.

- Key variables were engineered into numerical formats suitable for modelling
- Extracted relevant features and explored trends

2. Model Building & Evaluation

Processed data was analysed using Python to train and compare multiple ML models.

- Tested models: Linear Regression, Polynomial Regression with Ridge, K-Nearest Neighbor Regressor
- Selected best model based on prediction accuracy

3. Application Development

The best-performing models were deployed into a user-friendly interface.

- Frontend: Interactive dashboard for visual insights
- Backend: Model execution and predictions (Streamlit)

5. SYSTEMS DESIGN AND IMPLEMENTATION

PHASE 1. PREDICTIVE MODEL DEVELOPMENT AND BENCHMARKING

Feature Engineering and Exploratory Data Analysis

Number of features given in raw data - **6**
Number of features generated by team - **41**

Primary Output Variable = **Residual Days**
Residual Days = **Actual Delivery Date (ADD) - Estimated Delivery Date (EDD)**

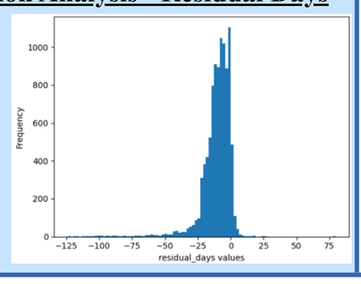
Log Transformation Analysis - Residual Days

Distribution

- Generally bell-shaped

Result:

- Log-transformation not required

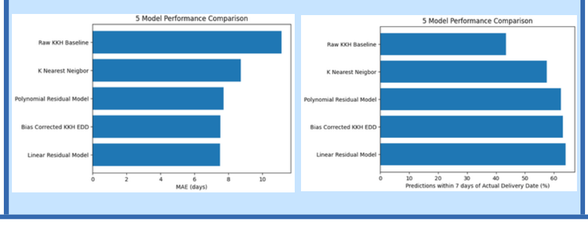


Candidate Model Development and Benchmarking

Number of features used for initial candidate models development - **11**

Model 1 - Bias Corrected KKH Baseline
Model 2 - Linear Residual (Regression)
Model 3 - Polynomial Ridge Residual (Regression)
Model 4 - K-Nearest Neighbor Regressor

Performance Comparision of Candidate Models



Evaluation of Different Feature Groups for Statistical Significance

Best performing model chosen from previous step for fine tuning - **Linear Model**

The central objectives of this process were:

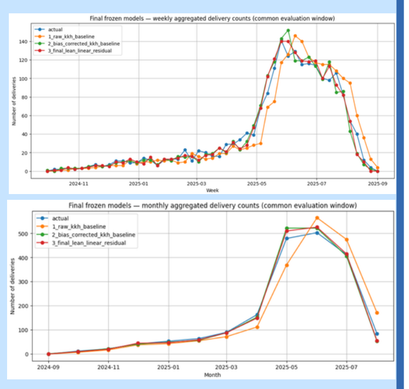
- To remove unwanted noise and extra complexity from the final prediction model
- Adopt leaner model with only relevant features
- Analyze the significance of the below feature groups within the original 11 features

Clinic Description | **Gestation Age Group** | **Appointment Calendar** | **EDD Calendar**

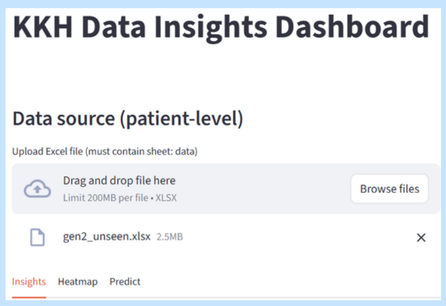
Design, Development and Freezing of Final Lean Linear Residual Model

Final number of features retained in the final **Linear Residual Model - 4**

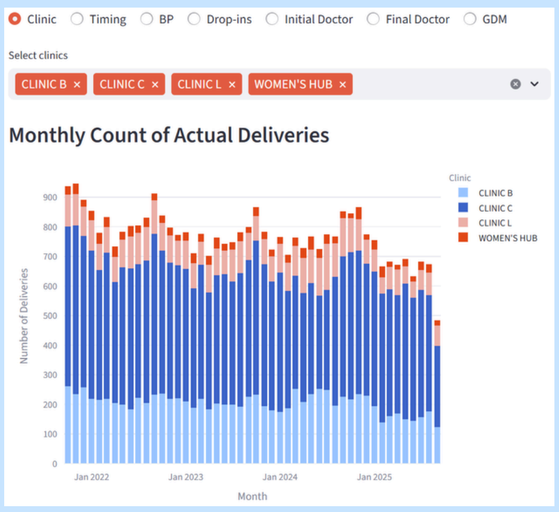
Performance comparison of **actual vs predicted** aggregate count of **weekly and monthly** deliveries.



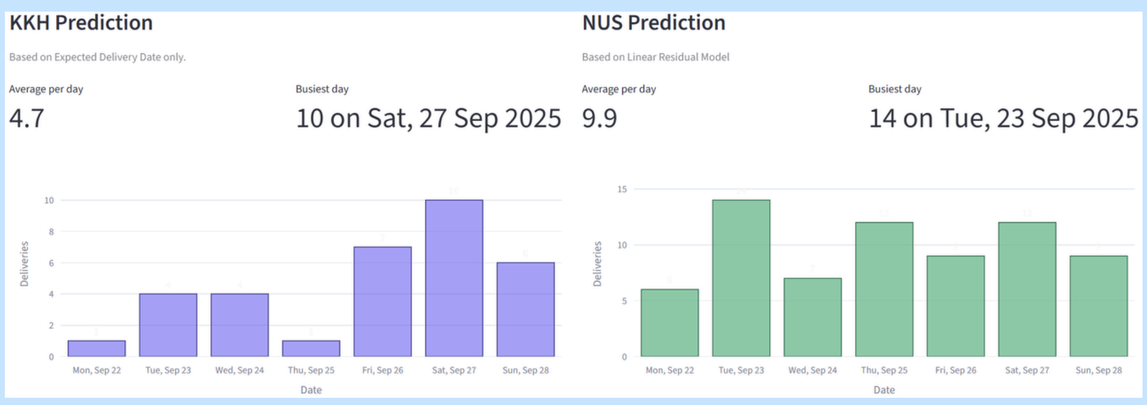
PHASE 2. STREAMLIT WEB APPLICATION SOFTWARE DEVELOPMENT



Patient-level Excel data is uploaded directly into the dashboard, the system serves as the entry point for automated analysis, visualisation and prediction.



Interactive filters allow users to explore delivery patterns by clinic, term classification, risk factors (Blood Pressure - BP and Gestation Diabetes Mellitus - GDM), and primary doctor.



Generates 7-day delivery forecasts, comparing KKH baseline with our model, key outputs such as average daily deliveries and busiest day support staffing and resource planning.

LIMITATIONS

Lack of Additional Features for Training:
Unavailability of actual data for antenatal patients diagnosed with GDM and high BP

Limited Scope for More Advanced Machine Learning (ML) Models:
Advanced and complex ML models such as Gradient Boosting and Tree based models require a larger feature space for enhanced generalization and learning over time

FUTURE DIRECTIONS

Data Enrichment:
Integrate additional antenatal and clinical features such as BP, GDM, and longitudinal patient history to improve prediction accuracy

Model Refinement & Validation:
Further validation on future patient cohorts through prospective testing, and periodic retraining and recalibration to maintain model robustness

SKILLS ACCQUIRED

1. Data Processing & Cleaning
2. Feature Engineering
3. Data Science & Machine Learning
4. Web Application Development