A DATA-DRIVEN APPROACH TO HEALTHCARE-ACQUIRED INFECTION CONTROL: PREDICTIVE MODELLING OF DISEASE TRANSMISSION USING RTLS DATA AND GRAPH NEURAL NETWORKS

Rivyesch Ranjan

INTRODUCTION

HAIs represent 21% of hospital capacity and contribute to 28,500 deaths annually in NHS England. The financial implications are substantial, costing the UK NHS an estimated £2.7 billion per year. The World Health Organization (WHO) defines HAIs as infections acquired during healthcare delivery, not present upon admission. These infections significantly burden healthcare systems and endanger patient well-being.

Robust surveillance systems are foundational in HAI control, but they can be labor-intensive and time-consuming. The advent of Information Technologies (IT) offers opportunities to automate surveillance and integrate Artificial Intelligence (AI) and machine learning.

METHODOLOGY

We employed an SIR model for epidemic simulation. Initially, a select few individuals were set as infected. Disease transmission was simulated using specific formulas. Three scenarios were created, each with varying infection rates: 50%, 60%, and 75% of the population.

The graph data is constructed using essential components: nodes representing current infection statuses, edge indices indicating contacts between individuals, edge attributes denoting proximity, and target labels for future infection statuses.

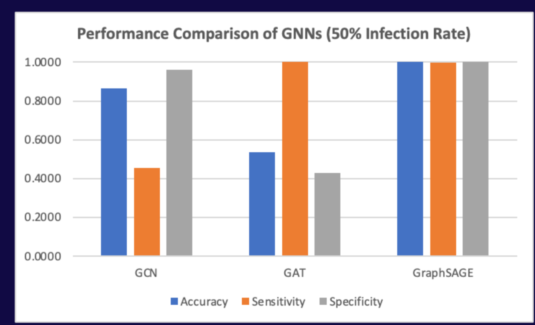
Each graph represents one timestep, and these graphs are sequentially compiled to form the temporal graph dataset.

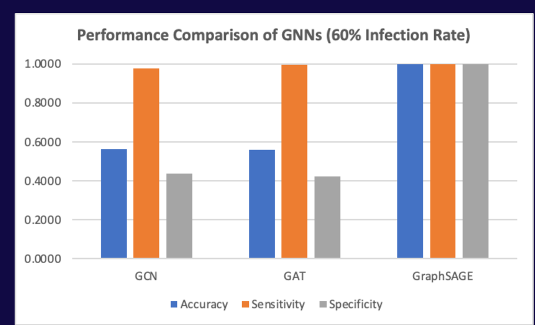
We utilized the temporal graph dataset to train and test three GNN model architectures (GCN, GAT, GraphSAGE). We employed a test set comprising one day's worth of graphs, with the remainder forming the training set. Hyperparameter tuning was conducted to achieve optimal GNN performance.

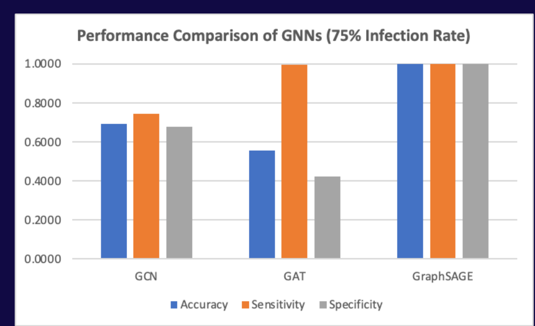
OBJECTIVE

Enhancing infection control in healthcare through advanced data integration and machine learning.

ANALYSIS





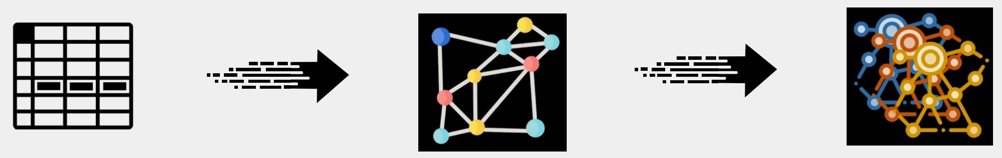


FUTURE WORKS

For future work, we aim to validate our models with real labels, explore diverse forecast horizons and rolling windows, integrate health information as node features, and further enhance hyperparameter tuning techniques

DATA

Proxximos provided pairwise proximity data obtained over a period of 3 weeks. Each timestep corresponds to a 5-minute interval. The data covered a population of 468 individuals. The SQL database table consists of five columns: position id, person id 1, person id 2, distance and timestep.



RESULTS/FINDINGS

- **GNN architectures exhibited varying performance, with GraphSAGE outperforming the other.**
- **The primary challenge lies in predicting infected individuals, as GNNs tend to miss some.**
- **Hyperparameter tuning revealed that incorporating a loss criterion with a positive weight improved sensitivity.**
- **Sensitivity analysis indicated that the model's performance varies depending on the percentage of the population infected.**

CONCLUSION

The research demonstrates the potential of Graph Neural Networks (GNNs) in predicting and controlling HAIs. The integration of RTLS data offers a promising avenue to enhance infection control practices in healthcare settings, minimizing the reliance on labor-intensive and costly traditional measures. By leveraging advanced data integration techniques and machine learning, we are moving closer to a future where HAI surveillance and response efforts are more efficient and effective, ultimately improving patient safety