White Paper: Optimizing Homeless Shelter Operations in Toronto Through Machine Learning

Executive Summary

Homelessness is a challenge in Canada, affecting thousands nightly and costing society billions annually. The City of Toronto provides services such as shelters, respite sites, drop-in programs, and warming centres to assist them. Recently, there has been unprecedented pressure on the shelter system due to the increased cost of living and lack of affordable housing. Traditional methods for managing shelter demands have often fallen short, especially during critical times of increased need. This white paper proposes a novel approach using a machine learning model to enhance the efficiency and responsiveness of homeless shelter operations. The proposed Long Short-Term Memory (LSTM) model aims to predict shelter demands with high accuracy, allowing for better resource allocation and potentially reducing the number of individuals turned away due to capacity issues. Preliminary results from this research indicate that the novel model outperforms traditional forecasting methods, offering a promising solution to a complex problem. This document outlines the development and implementation of this model, discusses its successes, and explores potential future enhancements and applications.

Introduction

Context and Importance of the Issue

In Toronto, over 9000 people are homeless each night due to various factors like family violence, mental illness, and job loss. The city provides services such as shelters, respite sites, drop-in programs, and warming centres to assist them. Recently, there has been unprecedented pressure on the shelter system due to the increased cost of living and lack of affordable housing. Despite having one of the largest shelter systems in the country, Toronto shelters had to turn away almost 200 people each day in December. Homelessness support services in Toronto encounter difficulties in forecasting the demand for shelter beds, resources, and staff. The current reactive approach relies on outreach efforts and community support when shortages arise. This strain on the shelter system highlights the need for improved demand forecasting and resource allocation to improve operation efficiency.

By enhancing forecasting and resource allocation for homeless support services, we can alleviate the strain on the shelter system and better meet the needs of those experiencing homelessness. This initiative aims to not only improve operational efficiency but also uphold the dignity and rights of all individuals in Toronto by ensuring access to stable housing, healthcare, and support services. Ultimately, by tackling homelessness, we strive to create a more compassionate and inclusive city where everyone has the opportunity to thrive.

Objective

The objective of this white paper is to present a transformative approach to managing homeless shelter operations through the application of advanced machine learning techniques. Specifically, it details the development and implementation of an LSTM model tailored to predict the fluctuating demands on shelters caused by various dynamic urban challenges. This document will guide policymakers, shelter managers, and stakeholders through the methodology, implementation strategy, and impact of this model, providing a roadmap for its adoption and potential expansion.

Problem Statement

To enhance the efficiency of Toronto's homeless shelter operations through the development of a machine learning model designed to accurately anticipate fluctuating shelter demands driven by dynamic urban challenges.

Review of Baseline Solutions

Forecasting models have been integral to managing shelter operations by predicting future demands based on historical data. The following are three commonly used baseline models in the field:

- **Autoregressive Models (AR):** These models have been widely used for time series forecasting due to their simplicity and effectiveness in certain scenarios. An autoregressive model predicts future behavior as a linear combination of past values. While AR models are useful for data with strong temporal correlations, they often fall short in scenarios where patterns change over time due to external factors, such as sudden economic shifts or weather-related impacts, which are typical in shelter demand forecasting.
- **Random Forest Regression (RFR):** This model uses an ensemble learning method for regression tasks, which involves constructing a multitude of decision trees during training and outputting the average prediction of the individual trees. Random Forest can capture nonlinear relationships better than AR models and is robust against overfitting. However, its predictive performance can degrade when encountering data outside the range of the training set, making it less effective in managing the variability and sudden changes often seen in shelter usage data.
- **Prophet Model:** Developed by Facebook, the Prophet model is designed for forecasting time series data that exhibits seasonal trends and patterns. It is particularly well-suited to data with clear and consistent patterns over time and can accommodate seasonality with multiple frequencies. While Prophet is powerful for handling daily and seasonal fluctuations in data, its performance

may diminish in the face of irregular events or when the data includes anomalies, which are common challenges in the context of homelessness.

Limitations of Baseline Models

Each of these models brings valuable insights and predictive power to different forecasting problems, but they have limitations when applied to the complex and dynamic environment of homeless shelter management:

- **Sensitivity to External Factors:** None of the baseline models inherently account for external influences without significant manual adjustments or additional data engineering, which can be cumbersome and inefficient in real-time operational settings.
- **Adaptability to Rapid Changes:** Homeless populations are affected by a variety of fast-changing factors that these models may not quickly adapt to, such as sudden policy changes or emergency situations.
- **Complexity of Integration:** Integrating these models into existing operational workflows can be challenging, as they may require extensive customization to align with specific operational needs and data environments.

Due to these limitations, there was a clear need for a more robust forecasting approach that could more accurately predict fluctuating demands and integrate seamlessly with shelter operations. This led to the development of the LSTM model, which aims to overcome these challenges by leveraging deep learning to enhance prediction accuracy and adaptability.

Proposed Solution

The machine learning solution for forecasting homeless shelter demand in Toronto leverages a Long Short-Term Memory (LSTM) model, which is particularly suited for time series data due to its ability to remember information over extended periods. This characteristic is crucial in understanding and predicting patterns in shelter usage that are influenced by complex, dynamic factors.

- **Data Collection:** We compiled an extensive set of data sources, including daily shelter occupancy rates, Government of Canada historical climate data, socioeconomic indicators like the New Housing Price Index, and emergency service call data. This rich dataset provided a comprehensive foundation for developing accurate predictive models.
- **Clustering and Geospatial Analysis:** To address the diverse nature of the shelter locations and their demand patterns, we applied clustering algorithms. This step categorized shelters into groups based on their geographical locations and other relevant features, allowing the model to make specialized predictions for each cluster rather than a one-size-fits-all forecast.

- **Initial Demand Forecasting:** Using the clustered groups, we trained the LSTM model on historical data to establish baseline demand forecasts. This model considered both regular patterns and seasonal fluctuations in shelter use.
- **Forecast Adjustment:** To refine the predictions, we incorporated a feature-based adjustment process. This involved analyzing additional data features that emerged from correlation analysis, such as specific events or sudden economic changes, to adjust the forecast in real-time, enhancing the model's responsiveness and accuracy.

Implementation Strategy

The implementation of the LSTM model is planned as a phased rollout, ensuring that the system can be refined and adjusted based on real-world feedback and performance.

- **Pilot Testing:** Initially, the model will be deployed in a select number of shelters. These pilot sites will be chosen based on their geographic and operational diversity to ensure that the model's effectiveness across different settings can be thoroughly evaluated. During this phase, the focus will be on closely monitoring the model's performance, gathering feedback from shelter managers, and making necessary adjustments.
- **Evaluation and Refinement:** After the pilot phase, we will conduct a detailed analysis of the model's accuracy and operational impact. This evaluation will include assessing how well the model's forecasts matched actual demand and identifying any issues in the data integration or model assumptions. Refinements will be made based on this analysis to improve the model before wider deployment.
- **Full-Scale Deployment:** Upon successful refinement and validation of the model, a full-scale implementation will be initiated across all shelters in Toronto. This broader deployment will be conducted in stages, gradually expanding the number of shelters using the model to ensure that the transition is smooth and that each shelter's specific needs and challenges are addressed.
- **Integration with Existing Systems:** A critical component of the deployment strategy is the integration of the LSTM model with existing shelter management systems. This process will involve developing interfaces or APIs that allow the model to seamlessly communicate with current software, ensuring that the predictive data is easily accessible and actionable for shelter managers. Collaborations with IT teams and software providers will be essential to facilitate this integration.
- **Training and Support:** To maximize the effectiveness of the new system, comprehensive training sessions will be provided for shelter staff. These sessions will cover how to interpret the model's forecasts, incorporate predictions into daily operations, and provide feedback for continuous

improvement. Ongoing support will also be available to address any issues that arise during daily use.

Implications and Future Work

The successful implementation of the LSTM model has several implications for policy, operational practices, and future research:

- **Community Impact Analysis:** Future work will involve detailed studies to measure how improvements in shelter forecasting affect the overall well-being of the homeless population. This includes tracking long-term outcomes such as duration of homelessness and success rates in finding permanent housing.
- **Geographic Expansion:** If the project achieves success in Toronto, a major area of future work is to adapt and apply the model to other cities. Each new implementation will help refine the model further, making it more versatile and effective.
- **Real-Time Implementation:** Ongoing development efforts should be focused on integrating real-time data feeds, allowing the model to adjust forecasts instantaneously as new data becomes available. This capability will be particularly valuable during sudden crises, such as extreme weather events or economic downturns.

Conclusion

The Adjusted Geospatial Multivariate LSTM model represents a potential significant advance in the use of technology to address social challenges. By providing precise and timely forecasts of shelter demand, the model not only improves operational efficiency but also contributes to a more humane and responsive shelter system. The success of this project highlights the potential of machine learning to transform public service operations, paving the way for broader applications in other areas of social need. This project exemplifies how innovative data-driven solutions can directly enhance the quality of life for vulnerable populations and provides a blueprint for future initiatives aiming to leverage technology for social good.