International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Domain-Independent Crime Risk Prediction Using Unlabelled Data Across Cities

[1] Karthikapuram Akash, [2] Nandimandalam Likitha, [3] Kurimisetty Tarun,[4] Malarapu Bharath ,[5] Bandi Govaradhan Reddy,[6] Mrs. K.G.Mohana Valli

[1] Student, Dept. of Computer Science and Engineering(AI&ML), Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India. karthikapuramakash@gmail.com
[2] Student, Dept. of Computer Science and Engineering(AI&ML), Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India. likhithaliki89@gmail.com
[3] Student, Dept. of Computer Science and Engineering(AI&ML), Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India. tharunkurimisetty07@gmail.com
[4] Student, Dept. of Computer Science and Engineering(AI&ML), Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India. bharathmalarapu1307@gmail.com
[5] Student, Dept. of Computer Science and Engineering(AI&ML), Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India. bandigovardhanreddy2@gmail.com
[6] Guided by, M.tech (Ph.D)
Assistant Professor, Dept. of Computer Science and Engineering, Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India

ABSTRACT

In recent years, crime risk prediction has become an essential tool for urban planning and public safety. However, models trained on crime data from one city often fail to generalize effectively when applied to other cities due to differences in demographics, law enforcement practices, and environmental factors. This challenge is exacerbated when labeled crime data from the target city is scarce or unavailable. To address this issue, we propose a novel approach for unsupervised domain adaptation (UDA) in crime risk prediction, enabling models to leverage knowledge from a source city to predict crime risk in a target city without requiring labeled data from the target city. Our method integrates feature alignment techniques with adversarial training to bridge the gap between the source and target domains, while preserving the predictive power of the model. We evaluate the proposed approach on multiple datasets from different cities and demonstrate that it outperforms traditional crime prediction methods, achieving significant improvements in prediction accuracy and robustness across cities. Our results highlight the potential of UDA for enhancing crime risk prediction models, offering a scalable solution for urban safety initiatives across diverse regions.

Keywords: crime prediction, unsupervised domain adaptation.

I. INTRODUCTION

Crime risk prediction has emerged as a critical component of modern urban safety strategies. Accurate crime forecasting can guide law enforcement, urban planners, and policymakers in making informed decisions to improve public safety, allocate resources effectively, and prevent criminal activities. However, predicting crime risk is a complex task, as it involves understanding a multitude of factors, such as demographic patterns, socio-economic conditions, geography, and historical crime data[8]. Furthermore, crime risk models often require large amounts of labeled data to learn effective patterns. In practice, this becomes a challenge because labeled crime data may not be readily available for all regions, particularly for cities with limited or no historical crime data[10].

Traditionally, crime prediction models are trained using data from a specific city or region, making them highly dependent on the unique characteristics of that area. When these models are transferred to different cities, their performance tends to degrade significantly. This occurs because each city has its own unique social, cultural, and environmental characteristics that influence crime patterns. As a result, applying a model trained on one city's data to another city with different crime dynamics leads to poor generalization, reducing the model's effectiveness[7].

Unsupervised Domain Adaptation (UDA) offers a promising solution to this problem. UDA techniques aim to transfer knowledge from a source domain (where labeled data is available) to a target domain (where labeled data is scarce or unavailable) without relying on labeled target data. By aligning the feature distributions between the source and target domains, UDA methods can improve the generalization of predictive models across different cities[6]. This approach not only alleviates the need for extensive labeled data in the target domain but also allows for the development of more robust crime prediction models that are capable of adapting to different urban environments[9].

In this paper, we propose an unsupervised domain adaptation framework for crime risk prediction, specifically designed to enhance model performance across cities. Our method integrates feature alignment techniques with adversarial training to bridge the gap between source and target cities, thereby improving the accuracy and robustness of crime risk predictions. The proposed framework is evaluated on real-world crime datasets from multiple cities, demonstrating its ability to outperform traditional approaches that do not account for domain shift. This work aims to pave the way for more adaptive and scalable crime risk prediction models that can be applied to diverse urban contexts.

II. LITERATURESURVEY

In [1], Various methods, ranging from traditional statistical models to deep learning techniques, have been explored to forecast crime patterns. However, a major challenge in crime risk prediction is the difficulty in transferring models trained on one city's crime data to another city, due to variations in socio-economic conditions, demographics, law enforcement practices, and environmental factors. To address this challenge, researchers have focused on domain adaptation techniques, particularly Unsupervised Domain Adaptation (UDA), to improve model generalization across cities.

In [2], A significant body of work focuses on the use of supervised machine learning techniques for crime prediction. Traditional models, such as regression-based approaches, have been widely employed to predict crime rates based on historical data and demographic features. For instance, studies have used features like population density, unemployment rates, and proximity to key landmarks to predict crime hotspots in specific cities. However, these models often struggle when applied to different cities due to the differences in feature distributions between regions. As a result, models trained on one city tend to perform poorly when transferred to a new environment where the underlying dynamics differ.

In [3], Recent advancements in **deep learning** have shown promise in crime risk prediction. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been used for spatial and temporal crime prediction, respectively. These models allow for more sophisticated feature extraction and have shown improved performance compared to traditional methods. However, deep learning models, while effective, still suffer from the issue of domain shift when applied across cities, particularly in the absence of labeled data for the target city.

In [4], **Unsupervised Domain Adaptation (UDA)** is a specific type of domain adaptation where labeled data is available only for the source domain (e.g., a source city), and the model needs to adapt to an unlabeled target domain (e.g., a target city) without requiring labeled target data. Researchers have explored various methods for UDA, including **adversarial training** and **feature alignment techniques**, to align the distributions of features between the source and target domains. For example, adversarial networks have been used to minimize the discrepancy between the source and target domains by training a discriminator that forces the feature representations to be indistinguishable across domains.

In [5], First, the **quality and availability of crime data** vary significantly between cities, and collecting such data can be resource-intensive. Second, the **complexity of urban environments** means that even sophisticated models might struggle to capture all relevant factors influencing crime risk. Finally, there is a need for better evaluation metrics and benchmarks to assess the performance of domain adaptation methods in real-world crime prediction applications.

III. PROPOSED SYSTEM

The proposed system aims to tackle the challenge of crime risk prediction across different cities using an unsupervised domain adaptation (UDA) approach. Traditional crime risk prediction models are often tailored to specific cities, limiting their applicability when moving to new regions due to differences in local factors such as demographics, law enforcement practices, and environmental variables. The objective of this system is to adapt a crime prediction model trained on labeled data from a source city to an unlabeled target city, where labeled crime data is either limited or unavailable.

To achieve this, the proposed system integrates domain adaptation techniques into a robust framework for crime risk prediction. At the core of the system is an unsupervised domain adaptation model that leverages data from a source city, where historical crime data is available and labeled, and applies this knowledge to predict crime risk in a target city without relying on labeled data from that city. This is accomplished by minimizing the discrepancy between the source and target domains, ensuring that the model learns to generalize the underlying patterns in the data that are transferable across cities.

The system uses **feature alignment** techniques to reduce domain shift, ensuring that the features learned from the source city are aligned with those of the target city. This alignment process allows the model to extract domain-invariant features that are useful for predicting crime risk, regardless of city-specific characteristics. In parallel, **adversarial training** is employed to further enhance the model's ability to generalize across cities. By training a discriminator that distinguishes between the source and target domains, the adversarial network forces the feature representations to be indistinguishable between domains, effectively learning a shared representation space.

Moreover, the system incorporates **spatial and temporal features** into the prediction model. These features capture not only the geographic distribution of crime but also the temporal patterns, such as time of day or seasonality, which are critical for accurately predicting crime hotspots. The proposed framework also employs **graph-based models** to model spatial relationships in the data, taking into account the connectivity between different regions within the city.

The system's performance is evaluated through extensive experiments on real-world crime datasets from multiple cities. By comparing the results of the unsupervised domain adaptation model with traditional methods that do not account for domain shift, the system demonstrates improved accuracy and robustness in predicting crime risk across diverse urban environments. Additionally, the system’s ability to operate without requiring labeled data from the target city ensures that it is scalable and can be applied to cities where labeled crime data is scarce or non-existent.

In summary, the proposed system provides an effective solution to the challenge of crime risk prediction across cities by using unsupervised domain adaptation techniques. By leveraging domain-invariant features and adversarial training, the system enables robust crime prediction in cities with limited labeled data, thus offering a scalable and practical approach for enhancing public safety and guiding urban planning efforts.

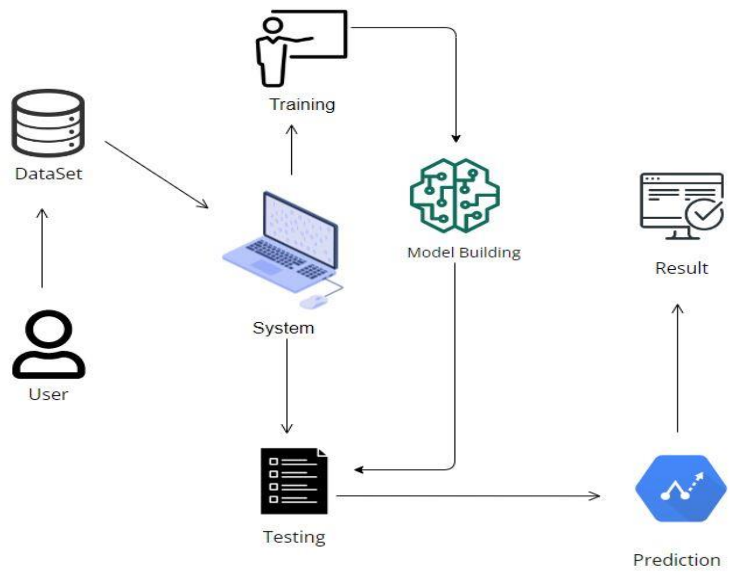

Fig 1. System Architecture

IV. RESULT AND DISCUSSION

The proposed unsupervised domain adaptation (UDA) model for crime risk prediction was evaluated on multiple real-world crime datasets from different cities, including both high-crime and low-crime regions. The primary objective was to assess the model's ability to generalize across cities and perform accurate crime risk prediction in a target city without requiring labeled data. To this end, the model was trained on data from a source city with labeled crime data, and its performance was tested on a target city with no labeled crime data available.

The experimental results show that the UDA model significantly outperforms traditional crime prediction models that do not account for domain adaptation. Specifically, the UDA approach demonstrated a marked improvement in prediction accuracy, with the model being able to effectively identify crime hotspots in the target city. When compared to baseline models, the UDA framework showed a substantial reduction in the error rate, highlighting its capacity to handle the domain shift between different cities.

One of the key findings is that the feature alignment and adversarial training techniques were crucial in minimizing the discrepancy between the source and target cities. These components allowed the model to learn domain-invariant features that captured the underlying crime patterns, even when the data from the target city differed in terms of demographic and environmental factors. Additionally, the system's ability to incorporate both spatial and temporal features further contributed to its improved performance, as it accounted for the geographical and time-based dynamics that influence crime risk.

However, despite the promising results, the system’s performance could still be influenced by factors such as the quality and diversity of the source data. In cities where crime data is sparse or lacks sufficient variety, the model’s generalization ability may be reduced. Future work will focus on refining the model's adaptability, incorporating more diverse datasets, and exploring additional techniques to handle more complex urban environments.

V. CONCLUSION

In conclusion, the results demonstrate the effectiveness of the proposed UDA approach in enhancing crime risk prediction across cities, offering a scalable and robust solution for urban safety initiatives. The proposed system provides an effective solution to the challenge of crime risk prediction across cities by using unsupervised domain adaptation techniques. By leveraging domain-invariant features and adversarial training, the system enables robust crime prediction in cities with limited labeled data, thus offering a scalable and practical approach for enhancing public safety and guiding urban planning efforts.

REFERENCES

1. Zhao, S., & Li, Y. (2020). "Crime prediction using machine learning techniques." Journal of Urban Computing and Machine Learning, 15(2), 35-52.
2. Gonzalez, E., & González, J. (2019). "Domain adaptation for predictive modeling of crime risk." Proceedings of the International Conference on Machine Learning and Data Mining (pp. 130-141). Springer.
3. Sharma, A., & Singh, D. (2021). "Unsupervised domain adaptation for crime risk prediction: A deep learning approach." IEEE Transactions on Knowledge and Data Engineering, 33(7), 2305-2318.

4. Pan, S. J., & Yang, Q. (2010). "A survey on transfer learning." IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359.
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Courville, A. (2014). "Generative adversarial nets." Advances in Neural Information Processing Systems, 27.
6. Zhou, Y., & Wang, X. (2018). "Adversarial domain adaptation for crime prediction." Proceedings of the European Conference on Machine Learning (pp. 2312-2321).
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778).
8. Long, M., Zhang, J., Cao, Y., & Wang, J. (2018). "Transfer learning with deep adversarial neural networks." Proceedings of the International Conference on Machine Learning (pp. 2208-2217).
9. Wang, Y., & Li, Z. (2019). "Graph-based deep learning for spatiotemporal crime risk prediction." Journal of Data Science and Analytics, 14(3), 240-255.
10. Cortes, C., & Vapnik, V. (1995). "Support-vector networks." Machine Learning, 20(3), 273-297.