# diffIRM: A Diffusion-Augmented Invariant Risk Minimization Framework for Spatiotemporal Prediction over Graphs

**Abstract**

Spatiotemporal prediction over graphs is challenging because real-world data suffers from the Out-of-Distribution (OOD) generalization problem, where the test data comes from a different distribution with regard to the training one. To address this issue, Invariant Risk Minimization (IRM) has emerged as a promising approach for learning invariant representations across different environments. However, IRM tends to degenerate to Empirical Risk Minimization (ERM) when applied to deep learning models due to overfitting. To avoid such a degeneration, we propose a diffusion-aumented invariant risk minimization (diffIRM) framework, that integrates the diffusion model into the IRM framework. Specifically, diffIRM augments the spatiotemporal graph data (such as human mobility) using a conditional diffusion model, which takes in conditions (such as COVID-19 case rates and demographic features) and generates diverse training environments. Then, the augmented spatiotemporal graph data is used to train Graph Neural Networks with the penalty of IRM. Theoretical proof is provided that the training environments generated by the diffusion model can guarantee the performance of IRM.

# 1 Model Structure

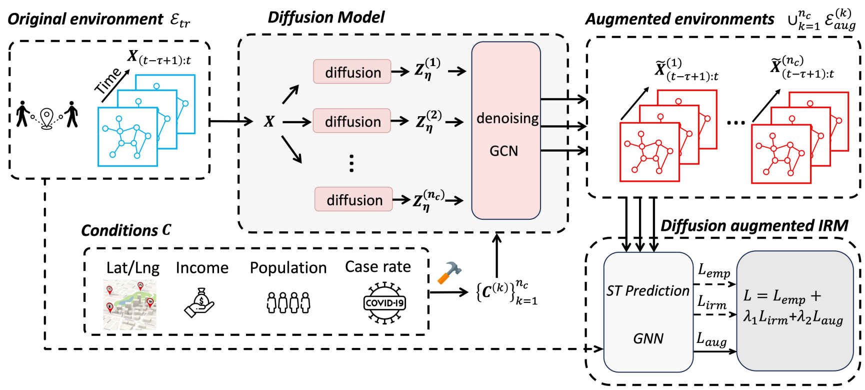



Figure 1: Model architecture of diffIRM

Fig. 1 provides an overview of our proposed methodology. We have introduced a method inspired by Denoising Diffusion Probabilistic Models[1] to construct a diffusion model capable of generating Urban Spatial-Temporal Data. As city ST data is graph based, and is already diffused (adding noise) to a Gaussian Noise by the forward diffusion process, we employed a Residual Gated Graph ConvNets[2] to denoise and reconstruct city ST data.

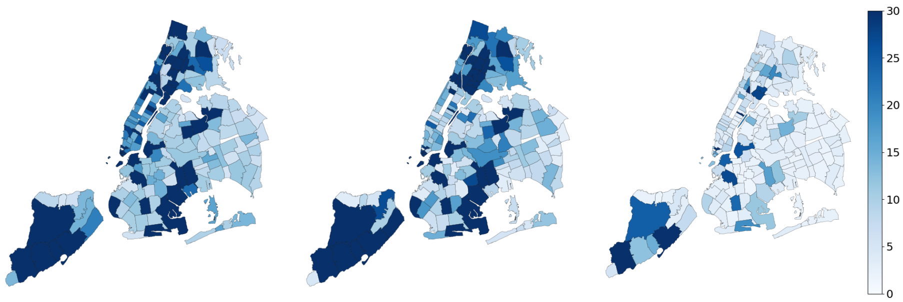
The diffusion model is pre-trained by minimizing the reconstruction error (MSE, KL divergence, ...) using the training dataset. Subsequently, this trained diffusion model is supplied with the original Spatial-Temporal (ST) graph data denoted as $X$ and a set of $n_c$ intervened conditions denoted as ${C^{(k)}}_{k=1}^{n_c}$, resulting in the generation of $n_c$ augmented environments. Both the original data within the initial environments and the augmented data are utilized as inputs for a ST prediction Graph Neural Network (GNN)[3].

Within this framework, the original data is employed to calculate two types of losses: the empirical loss ($\mathcal{L}emp$) and the Invariant Risk Minimization (IRM) loss ($\mathcal{L}irm$). Simultaneously, the augmented data contributes to the calculation of the augmented loss ($\mathcal{L}_{aug}$).

## 2 Experiment Results

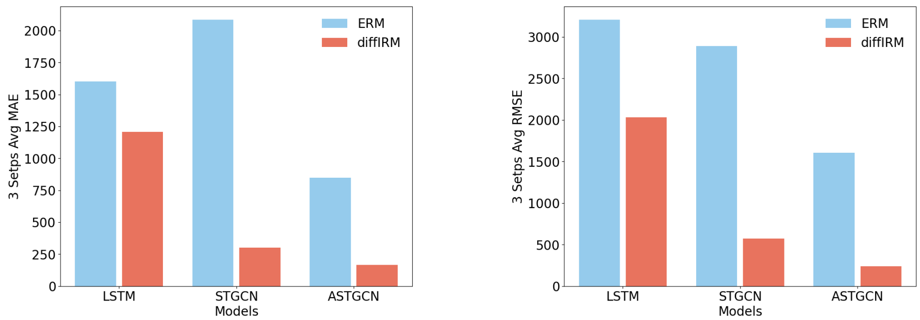
| | Safegraph | | | | PeMS04 | | | | PeMS08 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 step ahead | | 3 steps ahead | | 1 step ahead | | 3 steps ahead | | 1 step ahead | | 3 steps ahead | |
| | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| ERM | 828.0 | 1509.7 | 876.8 | 1604.8 | 327.8 | 610.9 | 355.1 | 722.8 | 124.4 | 1102.3 | 145.0 | 1260.4 |
| IRMv1 | 468.7 | 1076.8 | 518.0 | 1208.7 | 409.8 | 669.2 | 491.7 | 703.9 | 137.7 | 1333.3 | 145.1 | 1602.7 |
| REx | 434.2 | 957.2 | 1052.6 | 1955.8 | 427.4 | 703.2 | 485.3 | 693.8 | 145.6 | 1435.3 | 153.6 | 1836.5 |
| InvRat | 838.4 | 1596.4 | 1130.6 | 2095.3 | 409.8 | 669.2 | 491.7 | 703.9 | 162.4 | 1595.2 | 170.3 | 1957.3 |
| DiffAug | 221.1 | 438.6 | 254.6 | 524.4 | 109.2 | 277.9 | 163.9 | 282.6 | 117.9 | 907.0 | 120.4 | 959.3 |
| diffIRM (ours) | 122.4 | 292.3 | 132.4 | 308.6 | 81.9 | 273.2 | 81.9 | 263.8 | 84.6 | 806.4 | 94.6 | 912.7 |

Figure 2: Evaluation of different models using real-world mobility data




(a) ERM, 33.3% (b) IRM, 25.5% (c) diffIRM, 11.8%

Figure 3: MAPE heatmaps and overall MAPEs of different models for the Safegraph data.




(a) MAE

(b) RMSE

Figure 4: Ablation study of ST prediction GNN categories.

# References


[1] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

[2] Xavier Bresson and Thomas Laurent. Residual gated graph convnets. *arXiv preprint arXiv:1711.07553*, 2017.

[3] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 922–929, 2019.