Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.776658

MACHINE LEARNING PERSPECTIVE FOR ANALYSIS OF GEOSPATIAL DATA

Tanzeela Zarger[1] and Dr. Shobha Lal[2]
[1] Ph.D. Scholar Computer Science & Engineering, JVWU, Jaipur
[2] Professor of Mathematics & Computing, Department of Science andTechnology, JVWU, Jaipur

1. ABSTRACT
The characteristics of spatially explicit data are often inadequately handled in machine learning for spatial domains of application. At the same time, resources that can identify these properties and explore their impacts and how machine learning applications handle them are lagging behind. In this paper, we seek to identify and discuss the spatial properties of data that influence the performance of machine learning. We address existing research efforts and challenges in three main areas of machine learning: data analysis, deep learning and statistical inference. We will also discuss the existing end-to-end systems and highlight unresolved issues and challenges for future research in this area.
Keywords: Machine Learning, Machine Learning Algorithms, Spatial data, Geospatial Data, Spatial observation matrix, Classification.

2. Introduction:
Machine Learning (ML) is being used nowadays in almost every field to provide different solutions using structured and unstructured data. Machine Learning has proven its importance in different domains of application where spatial aspects are essential including land cover classification, cross-sectional characterization, urban growth, disaster management, transportation, and accident analysis, map visualization, delineation of geographic regions and habitat mapping, POI and region recommendation, trajectory and movement pattern prediction, point cloud classification, spatial interaction, spatial interpolation, and spatio-temporal prediction [1]. As spatial/Geospatial data has certain unique properties like spatial dependence, spatial heterogeneity, and scale, that makes it easy to design an effective Machine Learning Technique. This paper covers various Machine Learning methods for spatial data and aims to review some of the best recent practices of Machine Learning for spatial/Geospatial data.

3. Literature Review:
Pradhan [2] discussed the performance comparison of three well-known machine learning strategies: the decision tree (DT) approach, the support vector machine (SVM) algorithm, and the adaptive neural fuzzy inference system (ANFIS) method. The author reviewed aerial photographs and 113 landslide sites collected from field surveys. The study area focused on 340,608 pixels of which 8,403 pixels contain landslides. The generated dataset is divided into a training dataset that accounts for 50% of the data and the remaining 50% is considered a validation database. The processed images are sent to GIS technology to visualize the map representation of input parameters that are considered to be landslide sensitive data.

Reddy et al. [3] with the help of remote sensing and GIS technology, modeled the disturbance regimes as well as the biological richness in the area of the Similipal Biosphere Resource (SBR) located in the province of Orissa, India. The author calculated perturbation indices by considering various parameters causing perturbation such as proximity to the line, interlacing, fragmentation, segmentation, adjacency position, and porosity.

Lee et al. [4] using machine learning techniques, focused on removing as well as aggregating layers of buildings by classifying large-scale buildings into 0-discard, 1 retain, and 2 aggregate respectively. The author used a classification algorithm on the data to classify different buildings. Data were obtained by the authors from the National Institute of Geographic Information. The main classification algorithms used in this paper include decision trees, naïve, nearest neighbors, and support vector machines.

Lary et al. [5] highlight the essence of machine learning by solving problems found in GIS technology as well as remote sensing. Non-parametric analysis by regression and classification is introduced in the article to show the role of machine learning in improving the functionality of GIS and remote sensing applications. The author illustrated several clustering techniques and sources distributed over airborne particles and salt buoyancy data as case studies, respectively. The use of genetic programming in the field of GIS and remote sensing is illustrated through the various results presented in the paper.

Pourghasemi et al. [6] performed experiments using a support vector machine classifiers and GIS technology to generate a landslide susceptibility map at the town of Kalaleh located in the Golestan province of Iran. The kernel types used to classify and map the landslide susceptibility are as follows:
Linear Kernal Classifier,
Polynomial Kernal Classifier 2-degree
Polynomial Kernal Classifier 3-degree
Polynomial Kernal Classifier 4-degree
Radial Basis Function Kernel Classifier and Sigmoid Kernal Classifier.

Furlenello et al. [7] discussed the application of machine learning techniques in GIS techniques for dynamic analysis. The authors used GRASS tools (Geographic Resource Analysis Support System), the R programming language, and Postgre SQL technology to model and analyze the geospatial epidemiological data.

Fandino et al. [8] propose ways to analyze crime data using machine learning in the background. In addition, emphasis was placed on discovering crime samples using the R programming language. The paper adopted data mining strategies and performed a histogram analysis on a dataset.

Flaxman et al. [9] proposed a predictive solution for crime events in space-time. The paper used advanced techniques such as Kernel Hilbert Space's RKHS acronym reconstruction to approximate Gaussian Processes with auto-recovery smoothening Kernel. Furthermore, the proposed procedure focuses on improving the performance of two widely used areas of crime analysis which are the acronym (KDE) Kernel Density Estimation and the acronym for (SEPP) Self Exciting Pint Process.

Murayama [10] discussed the basic concepts of deep learning as a subfield of machine learning in the context of prediction and simulating urban growth, urban sprawl, and urban development. Many socioeconomic factors are selected to determine attributes such as market behavior, population, density, transport infrastructure, and government policies regarding land. Big data management is a challenge when machine learning is the main technology. Therefore, the author has typified the use of deep learning proficiency enough to handle spatial big data to perform predictive analysis on multi-dimensional geospatial big data.

Castelluccio et al. [11] examined the features of convolutional neural network (CNN) in the context of semantic classification on remote sensing classification. The authors reviewed two recent architectures, CafffeNet and GoogLeNet, for semantic classification using three different methods. This paper focuses on using pre-trained and fine-tuned networks for training rather than conventional training approaches to reduce design time and solve overfitting problems. The authors perform experiments on three remote sensing-based datasets to confirm their proposed framework.

Scott et al. [12] investigated the use of a deep convolution neural network (DCNN) for land cover classification using high-resolution data obtained from remote sensing. Two methods of convergence have been proposed namely transfer learning (TL) and new data augmentation techniques. The authors report that TL allows the DCNN to be started with well-preserved extraction and the data station improves the robustness of the DCNN. The authors considered adding the UCMerced dataset to realize their proposed solution in practice.

Bui et al. [13] analyzed the occurrence of malaria based on sociophysical factors in the Daknong region of Vietnam using remote sensing, GIS, and machine learning classification. The authors performed an accuracy assessment using receiver operating characteristics (ROC) and paired tests. This paper focuses on using assembly models and proves that the random subspace model is the best fit. The motivation behind the study was to perform vulnerability mapping on malaria data so that control measures could be effectively implemented based on the map.

Zhang et al. [14] propose a new model based on deep learning that performs predictive mapping based on sparse spatiotemporal events. This article used the acronym GLDNet closed local broadcast network to represent a graph of networked structured data. In addition, this article has attempted to distinguish between two mechanisms for spatial and temporal data, which were closed networks means temporal data, and GLDNet means spatial data.

Duan et al. [15] focus on feature extraction from crime datasets using deep neural networks. The proposed model, acronym STCN, refers to the SpatioTemporal Crime Network, which utilizes the essence of a CNN with a high-dimensional metaphysical crime dataset. Crime risk prediction model based on the provided dataset.

Rao et al. [16] introduced two main multimodel context learning techniques for modeling the correlation between visual images captured by an autonomous underwater vehicle and acoustic depth measurement data obtained before here by remote sensing. A multi-tier architecture is used to find the common distribution between visualization and depth measurement methods.

4. Methodology:

To conduct Machine learning of geospatial data, we need to add location, distance, or topological relations to the process of learning. The process can be categorized into two sections, the spatial observation matrix and the Machine learning algorithms.

1. **Spatial Observation Matrix**:

To include spatial characteristics in Machine Learning we have to find a representation for these characteristics in the observation Matrix. The principle is that we can use Machine Learning methods like SVM, Decision Tree, neural networks, etc. without making changes to the algorithms after we design and engineer the observation matrix to include spatial properties. Various aspects like spatial sampling, spatial features, dimensionality reduction, and handling of missing data are considered while creating a spatial observation matrix that is used as an input to the Machine Learning algorithm. These are discussed below.

➢ Spatial Sampling

Though there has been a lot of progress in the technology for the collection of spatial data there are still challenges faced in getting the optimized samples of data for training a Machine Learning model. The data sample set should represent the complete distribution or entire population from a statistical point of view. It is not only the scarcity of samples that leads to challenges for learning. Though the Learning process will not be affected by the sampling it may result in the overestimation of the accuracy of learning. So we should make sure that we should be careful in sampling the data that we feed into the algorithm. For a proper review of the spatial sampling methods. Inter-class imbalance, in which the number of samples in class categories is highly uneven, can also degrade the accuracy of data classification. If we have more data samples, the performance of a class will be higher, compared to the classes with fewer samples though it overestimates the overall accuracy.

➢ Spatial Features

Various methods exist to include the spatial components of data into the observation matrix. One way is to add spatial references directly to the data matrix as attributes. Practically there are two ways to implement it. One way is to add coordinates for all the observations along with semantic attributes to the observation matrix [17,18] and the other option is to add observations tied to a region to the observation matrix as fixed effects of that region [19,20]. This way is preferable for handling inclusion relationships. Though it cannot capture complex structures, it can capture geometric, spectral, textual, statistical, contextual, and relational entities apart from spatial reference information, and that can be created as new features and can be directly added to the observation matrix [1].

- **Dimensionality Reduction**

A huge number of input variables may end up in the observation data matrix as the result of the Machine Learning task. A large number of interrelated variables may impact learning in many ways. More training data may be needed and it will also increase the processing time because of the correlation between the variables [21]. There are different methods of feature selection. Dimensionality reduction methods are also a good solution to handle various useful variables, in particular when the influence of each variable is not of interest. The structure of the variance-covariance matrix needs to be understood to minimize unnecessary variables. The problem is that the calculation of the variance-covariance matrix $C_{n \times n}$ for a given observation matrix $X_{m \times n}$ is computationally expensive for a large set of variables. Many dimensionality reduction methods exist, including Principal components analysis, factor analysis, self-organizing maps, and independent components analysis.

- **Missing Data**

Though the data is easily available the data created by the other processes have gaps in spatial and temporal dimensions. So, missing data is a big challenge, and many analyses cannot be implemented unless this problem is dealt with. Missing values can be dependent on their neighboring points, or specific patterns [22]. There are various approaches to solving this problem, such as accumulating data at a coarser granularity, removing instances of missing values from the data set and input values. Although data transfer adds a pre-processing step to the analysis, it leverages the existing data and avoids loss of information due to aggregation, and discards some observations. To impute values for data sets with missing values, spatial prediction methods can be used. The well-known methods for spatial prediction are spatial statistical models like geographically weighted regression and geostatistical approaches such as kriging [23].

4.2. Learning Algorithm

Instead of processing the spatial features with traditional methods, we can feed these spatial properties into the existing Machine learning algorithm. The various Machine Learning techniques that have found attention in spatial science include decision trees, random forests, SVM, neural networks, and deep neural networks. Here we will discuss a few of these algorithms.

- **Decision Trees**

To overcome the violation of assumption i.i.d, a Decision tree is a popular Machine Learning method used for spatial problems. A spatial entropy decision tree classifier uses information gained along with spatial autocorrelation to select tests of candidate tree nodes in a raster spatial model. Hierarchies of clusters of similar data are identified with the help of PCT, which is a multi-task approach, and a predictive model is associated with each group. To maximize variance reduction within a cluster a test is run while considering splitting a group at a node. A term based on global measures of spatial autocorrelation was added to this test to account for spatial non-stationarity in the target variable. Salt and pepper noise is one of the common problems when classifying images using decision trees. This problem occurs when the predicted label of a particular pixel is different from the surrounding pixels. Focaltest-based spatial decision trees (FTSDT) use local indicators of spatial association-Lisa [24] as spatial autocorrelation statistics to measure spatial dependencies between adjacent pixels.

- **Support Vector Machines**

The support vector machine algorithm is mainly used for classification and regression problems [25]. The purpose of SVM is to map the original input space to a higher dimensionality feature space where the observations are separable by hyperplanes. The hyperplane that maximizes the margin width (e) is optimized among all possible hyperplanes [26]. Support Vector Machines performs good in high-dimensional spaces. It is powerful in generalization and less sensitive to class imbalance [26]. Researchers suggested an extension of SVM called support vector random field that uses conditional random field (CRF) to explicitly models spatial dependencies in the classification. It has two components: the observation-matching potential function and the local-consistency potential function. The observation-matching potential function models the relationship between the observations and the class labels using an SVM classifier, and the local-consistency potential function models the relationship to neighborhood labels.

- **Self-Organizing Maps**

Self-Organizing Maps (SOM) is one of the nonlinear clustering methods that has been used with spatial and non-spatial data [27].This is a simple neural network with no hidden layers. It maps an n dimensional feature vector to a regular grid of four or six neighbor neurons in the output layer, initialized with n weights. We first use a similarity measure to find more neurons that are similar input feature vector and then weights of the activated neurons and their neighboring neurons are adjusted to make them even more similar to the input vector. This process is repeated for the set of input feature vectors. Finally, this creates a spatial organization of neurons with different units far apart in one-, two-, or three-dimensional space.

- **Radial Basis Function Networks**

The two hidden layers in the RBF network are the output layer and an input layer. The input vector and the neurons weighted distance are computed in the RBF network using a radially symmetric activation function, which is usually Gaussian. Instead of the linear relationship between the input vector and the neurons in the hidden layer [28], we compared the RBF network with the MLP network for urban change, modeling and found that RBF provides higher prediction accuracy. The researchers used the RBF network for spatial interpolation, including a semivariogram model in which the hidden layer neurons are the centroids of the observations.

- **Adaptive Resonance Theory Networks**

Networks based on Adaptive Resonance Theory (ART) are a suite of neural networks used for spatial interaction flows, crop classification, and land-use change applications [29]. These networks are supervised, self-organizing, and self-stabilizing neural networks that can quickly learn in deviant environments [30]. The best-known ART-based network is Fuzzy ARTMAP, which combines ART-based networks with fuzzy logic [31]. It contains two input modules, Arta and Artb, each with two layers connected by a Map Fields module. Arta matches the input vector to the most similar neuron in the second layer. If the vector does not resemble the current neuron in memory, a new neuron is created. This property allows Arta neural networks to adaptively change network topology and add new experiences to memory. Artb, which supports class labels, is linked to Arta through the map module. However, Fuzzy ARTMAP can be considered a novel pattern as it depends on the quality of the training data and its sensitivity to noise and outliers.

- **Deep Graph Neural Networks**

CNN's have been used quite extensively for image classification and segmentation. However, many problems such as social and biological networks, cannot be represented in grid form, making it convolution difficult to apply. Thus, attempts were made by researchers to extend neural networks to phenomena that are best portrayed with graph structure. This is how graph neural networks (CNNs) were introduced. Recently growing attempts have been made to generalize convolution to graphs that can be categorized into spectral and non-spectral approaches [32]. Spectral methods create a spectral representation for the graph and apply convolution through the graph Fourier Transform [33]. The challenge with these types of graphs is, if the structure of the graph changes, a trained model of the old structure cannot be directly applied to the graph of the new structure. Non-spectral methods directly use convolution on the nearest neighbor in the graph [34]. This approach is relatively new and has shown impressive performance in many applications, such as disease spread forecasting [35], traffic analysis [36], medical diagnosis and analysis [37], and natural language processing [28]. The application of GNNs has yet to be explored in spatial domains especially for non-grid-based spatial data such as social networks.

5. **Conclusion:**

We examined the literature where machine learning intersects with spatial domains, where the data exhibits special properties such as spatial dependence, spatial heterogeneity, and scale. We have discussed two main approaches in this part of the document, which are spatial observation matrices and learning algorithms. The observation matrix explicitly deals with the spatial properties of the data before the learning process begins. In other words, no modification of the learning algorithm is made after this step. We also discuss that taking into account spatial properties in missing data processing and sampling strategies is essential for any spatial application of ML. In addition to these problems, the generation of new spatial features as one of the main approaches to augment the observed matrix with new spatial properties of the data has been discussed. To date, much of the literature on machine learning from spatially explicit data has used spatial features mainly because the idea came naturally, because of extensive research in geospatial information science. Management has focused on these issues for the past two decades, and because of this approach allows existing ML algorithms to be used without further modification. Many of these methods have been used successfully for a variety of applications ranging from point cloud classification to orbital analysis and pattern recognition in satellite imagery. We also discussed how spatial properties can be explicitly handled in another component of ML, namely learning algorithms, an approach that has only recently begun to be explored. Here, the spatial properties are resolved in the representation of the learning algorithm or the objective function rather than at the level of the observation matrix**.**

More study is required to learn more about the spatio-temporal domains. Simultaneous learning across space, time, scale, and hierarchy is introduced by Deep neural networks that are based on a combination of LSTM and CNN. When supplemented with reinforcement learning to add feedback in systems, which is the case in many spatial, social and environmental applications, they can fulfill the dream of a single universal Machine Learning method [27]. But deep neural networks have also various limitations like a huge amount of training data required and also they have a huge number of parameters that makes them expensive in terms of the computation. Also, the complexity of the DNNs is beyond the limit because of the arbitrary nature of their architectural design. In addition to the above suggestions for future research, to analyze geospatial data, a long-term research path in this area is needed.

References:

[1] Behnam Nikparvar and Jean-Claude Thill -Machine Learning of spatial data

[2] B. Pradhan, A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using gis, Computers & Geosciences 51 (2013) 350–365.

[3] C. S. Reddy, C. Pattanaik, M. Murthy, P. Roy, Spatial modeling for biolog985 ical richness analysis in similipal biosphere reserve, orissa, india, Journal of Bioresource Conservation and Management (2008) 285–298

[4] J. Lee, H. Jang, J. Yang, K. Yu, Machine learning classification of buildings for map generalization, ISPRS International Journal of Geo-Information 6 (10) (2017) 309.

[5] D. J. Lary, A. H. Alavi, A. H. Gandomi, A. L. Walker, Machine learning in geosciences and remote sensing, Geoscience Frontiers 7 (1) (2016) 3–10.

[6] H. R. Pourghasemi, A. G. Jirandeh, B. Pradhan, C. Xu, C. Gokceoglu, 1000 Landslide susceptibility mapping using support vector machine and gis at the golestan province, iran, Journal of Earth System Science 122 (2) (2013) 349–369.

[7] C. Furlanello, M. Neteler, S. Merler, S. Menegon, S. Fontanari, A. Donini, A. Rizzoli, C. Chemini, Gis and the random forest predictor: Integration in r for tick-borne disease risk assessment, in: Proceedings of DSC, 2003, p. 2.

[8] P. M. Fandino, J. B. Tan Jr, Crime analytics: Exploring analysis of crimes through r programming language, Science 132 (2019) 696–705

[9] Flaxman, M. Chirico, P. Pereira, C. Loeffler, et al., Scalable highresolution forecasting of sparse spatiotemporal events with kernel methods: 1015 a winning solution to the nij "real-time crime forecasting challenge", The Annals of Applied Statistics 13 (4) (2019) 2564–2585.

[10] I.-P. D. Y. Murayama, T. A.-G. Hao, Machine learning in geoscience.

[11] M. Castelluccio, G. Poggi, C. Sansone, L. Verdoliva, Training convolutional neural networks for semantic classification of remote sensing imagery, in: 1020 2017 Joint Urban Remote Sensing Event (JURSE), IEEE, 2017, pp. 1–4.

[12] G. J. Scott, M. R. England, W. A. Starms, R. A. Marcum, C. H. Davis, Training deep convolutional neural networks for land–cover classification of high-resolution imagery, IEEE Geoscience and Remote Sensing Letters 14 (4) (2017) 549–553.

[13]T. Bui, Q.-H. Nguyen, V. M. Pham, M. H. Pham, A. T. Tran, Understanding spatial variations of malaria in vietnam using remotely sensed data 1005 integrated into gis and machine learning classifiers, Geocarto International 34 (12) (2019) 1300–1314

[14] ng, T. Cheng, Graph deep learning model for network-based predictive hotspot mapping of sparse spatio-temporal events, Computers, Environment and Urban Systems 79 (2020) 101403.

[15]Duan, T. Hu, E. Cheng, J. Zhu, C. Gao, Deep convolutional neural net1035 works for spatiotemporal crime prediction, in: Proceedings of the International Conference on Information and Knowledge Engineering (IKE), The Steering Committee of The World Congress in Computer Science, Computer . . . , 2017, pp. 61–67.

[16] D. Rao, M. De Deuge, N. Nourani-Vatani, S. B. Williams, O. Pizarro, Multimodal learning and inference from visual and remotely sensed data, 1030 The International Journal of Robotics Research 36 (1) (2017) 24–43

[17] Martin, R.; Aler, R.; Valls, J.M.; Galván, I.M. Machine learning techniques for daily solar energy prediction and interpolation using numerical weather models. Concurrency and Computation: Practice and Experience 2016, 28, 1261–1274.

[18]Zanella, L.; Folkard, A.M.; Blackburn, G.A.; Carvalho, L.M. How well does random forest analysis model deforestation and forest fragmentation in the Brazilian Atlantic forest? Environmental and ecological statistics 2017, 24, 529–549

[19] Anselin, L.; Arribas-Bel, D. Spatial fixed effects and spatial dependence in a single cross-section. Papers in Regional Science 2013, 92, 3–17.

[20] Sommervoll, Å.; Sommervoll, D.E. Learning from man or machine: Spatial fixed effects in urban econometrics. Regional Science and Urban Economics 2019, 77, 239–252.

[21] Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. An object-based convolutional neural network (OCNN) for urban land use classification. Remote sensing of environment 2018, 216, 57–70.

[22] Qu, L.; Li, L.; Zhang, Y.; Hu, J. PPCA-based missing data imputation for traffic flow volume: A systematical approach. IEEE Transactions on intelligent transportation systems 2009, 10, 512–522

[23] Cressie, N. The origins of kriging. Mathematical geology 1990, 22, 239–252.

[24] Anselin, L. Local indicators of spatial association—LISA. Geographical analysis 1995, 27, 93–115

[25] Vapnik, V. The nature of statistical learning theory; Springer science & business media, 2013.

[26] Effati, M.; Thill, J.C.; Shabani, S. Geospatial and machine learning techniques for wicked social science problems: analysis of crash severity on a regional highway corridor. Journal of Geographical Systems 2015, 17, 107–135.

[27] Kohonen, T. Self-organization and associative memory; Vol. 8, Springer Science & Business Media, 2012.

[28] Shafizadeh-Moghadam, H.; Hagenauer, J.; Farajzadeh, M.; Helbich, M. Performance analysis of radial basis function networks and multi-layer perceptron networks in modeling urban change: a case study. International Journal of Geographical Information Science 2015, 29, 606–623.

[29] Gong, Z.; Thill, J.C.; Liu, W. ART-P-MAP neural networks modeling of land-use change: accounting for spatial heterogeneity and uncertainty. Geographical Analysis 2015, 47, 376–409

[30] Carpenter, G.A.; Grossberg, S.; Reynolds, J.H. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. Neural networks 1991, 4, 565–588.

[31] Carpenter, G.A.; Grossberg, S.; Markuzon, N.; Reynolds, J.H.; Rosen, D.B.; others. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE Transactions on neural networks 1992, 3, 698–713.

[32] Veliˇckovi´c, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. arXiv preprint arXiv:1710.10903 2017.

[33] Estrach, J.B.; Zaremba, W.; Szlam, A.; LeCun, Y. Spectral networks and deep locally connected networks on graphs. 2nd International Conference on Learning Representations, ICLR, 2014, Vol. 2014.

[34] Hamilton, W.L.; Ying, R.; Leskovec, J. Inductive representation learning on large graphs. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 1025–1035.

[35] Wu, D.; Gao, L.; Xiong, X.; Chinazzi, M.; Vespignani, A.; Ma, Y.A.; Yu, R. DeepGLEAM: a hybrid mechanistic and deep learning model for COVID-19 forecasting. arXiv preprint arXiv:2102.06684 2021.

[36] Ye, J.; Zhao, J.; Ye, K.; Xu, C. How to build a graph-based deep learning architecture in traffic domain: A survey. IEEE Transactions on Intelligent Transportation Systems 2020.

[37] Ahmedt-Aristizabal, D.; Armin, M.A.; Denman, S.; Fookes, C.; Petersson, L. Graph-Based Deep Learning for Medical Diagnosis and Analysis: Past, Present and Future. arXiv preprint arXiv:2105.13137 2021.

[38] Vashishth, S.; Yadati, N.; Talukdar, P. Graph-based deep learning in natural language processing. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD; ACM, 2020; pp. 371–372