INSTITUTION DETAILS

Province	Sindh	City	Karachi
Institution	National University of Computer and Emerging Sciences (FAST-NU)	Campus	Karachi
Department	Computer Science	Degree Level	BS
Degree Program	Computer Science	Telephone	
Fax	-		

SUPERVISOR DETAILS

Name	Dr. Murk Marvi	Gender	Female
Mobile	-	Office No	-
Email	Murk.marvi@nu.edu.pk	Designation	Assistant Professor
Qualification	Doctor of Philosophy(Ph.D.)		

Co-SUPERVISOR DETAILS

Name	Ms. Anam Qureshi	Gender	Male
Mobile	03337288578	Office No	Research Lab
Email	Anam.qureshi@nu.edu.pk	Designation	Lecturer
Qualification	-		

Head of Department Details

Name	Dr. Zulfiqar Memon	Mobile No.	-
Email	zulfiqar.memon@nu.edu.pk	Gender	Male

PROJECT DETAILS

Project Title	Development of a spatial and temporal based COVID-19 predictor for Pakistan
Group Details	1. Ovaiz Ali (18K-0137) 2. Zaeem Ahmed (18K-0166) 3. Rabil Maknojia (18K-1183)
Project Area of Specialization	Machine Learning & Deep Learning
Project Start Date	09-September-2021 **Project End Date** June-2022
Project Summary(Less than 2500 Characters)	Pakistan has seen a large-scale spread of coronavirus (COVID-19) with unusual patterns across the country. Starting in a few provinces, this severe disease spread down the Boot and is now found in virtually all of Pakistan's provinces. Differences in epidemic dynamics in Pakistan highlight the critical need for a strong national coordinating level for uniform implementation of control measures, as well as the relevance of local forecasting. Since the outbreak of COVID-19, one of the most often asked questions by public officials has been about determining the peak of this contagious disease. Analyses and forecasts in both space and time can provide a clear picture of which regions will be most affected and when. Such an analysis can provide decision-makers with enough time to intervene in local policy. Furthermore, the use of spatial-temporal-based predictive models at the provincial level can assist public decision-makers in better planning health policy actions by substantially improving forecasts of the number of infected persons.
Project Objectives(Less than 2500 Characters)	Following are the main objectives of this research: 1. To provide a processed and cleaned dataset on the COVID-19 pandemic's spatial and temporal trends in Pakistan. 2. To determine the influencing factors which can increase or decrease the rate of spread of COVID-19 in Pakistan. 3. To develop a machine learning-based predictor for forecasting the rate of COVID-19 spread in Pakistan.
Literature Review/Background Study	The COVID-19 virus spreads in a Spatial-temporal manner. The Spatial-temporal spread pattern was discovered utilising methods such as geographic clustering, hotspot identification, [1, 2] and the direction of virus being transmitted is using the spatial heterogeneity and homogeneity techniques [1]. Furthermore, approaches such as geographic statistical analysis, epicentre incidence rates, and the "space-specific strategy" are utilised to identify the strategies to minimise the virus's spatial spread [1]. Previous research discovered that the transmission of the COVID-19 virus was influenced by the following geographical factors. ● The same geographical pattern (high probability of a spatial pattern) as previous viruses such as influenza, pneumonia, respiratory Syncytial virus, malaria, dengue, zika, HIV, and Choles. [1. 2] ● The spread of infectious illnesses, particularly COVID-19, is also linked to human movement, according to the study. [1] ● Infectious illnesses are geographically specific. [1, 2]

	Geographic information systems (GIS) methodology prepares a great foundation for combining explicit data reflecting disease and their interpretation in relation with the population habitancy, health, social services environment, and nature. [1, 2]
	According to Tobler's first law, [1] It mainly focuses on, **where is the disease concentrating, and in which direction it is spreading?** in order to **determine the trend of any disease and spread over space.** In all the studies based on spatial-temporal analysis of COVID-19, the maps of respective countries are coloured on a specific scale on a weekly/monthly basis to determine the spread of the virus and different factors such as hotspots, epicentres, and cold spots are plotted which, plays an important role in the virus's spread[1, 2]. Moreover, it could be deduced that the COVID-19 cases are not only affected by climatic and air pollution factors but also influenced by geographical landscapes, economic conditions, demographic disease variations, genetic factors, health care system, number of testing and age differences, etc. [5] Furthermore, one major factor contributing to the spread of the COVID-19 is a large number of undiagnosed infected individuals.[9]
Project Implementation Method(Less than 2500 Characters)	To begin, we will gather data from various sources from the outbreak of COVID-19 in Pakistan to a predefined date. If necessary, we will use web scraping to obtain crucial data. This information is part of the COVID tracking project, which gathers COVID-19 statistics on test counts, cases, hospitalizations, and patient outcomes, among other things. Second, we will conduct data pre-processing and make use of a variety of libraries of python (like Pandas, Matplotlib, Seaborn, Scikit, Scipy, Beautifulsoup, Tensorflow, and Pytorch) that will be critical at this point. Then, to extract useful insights from the processed data, we will do statistical analysis. After that, we will use Machine Learning Algorithms to construct our predictive model. In the end, we will evaluate the performance of the developed predictor by using several performance metrics.
Benefits of the project(Less than 2500 Characters)	The value of this study is that it aids in determining the spatial and temporal patterns of the COVID-19 epidemic in Pakistan. Determining the spatial trends of COVID-19 fatalities and recoveries can also help with disease control and prevention. Through such work, it will be possible to determine which variables affect death and recovery. Furthermore, spatial and temporal analysis and forecasts show which regions will be most affected and when giving decision-makers the time to respond to local policies. The use of spatial-temporal models at the provincial level can help decision-makers in better planning healthcare policy initiatives by considerably boosting forecasts of the number of affected persons.
Technical Details of Final Deliverable (less than 2500 characters)	A predictor for forecasting the spread of COVID-19 spread in different regions of Pakistan.

Final Deliverable of the Project	A predictor for forecasting the rate of COVID-19 spread in different regions of Pakistan, along with the dataset based on the features which are highly corelated.
Core Industry	Health and Government Sector
Other Industries	-
Core Technology	Data Science
Other Technologies	Machine Learning, Statistics
Sustainable Development Goals	Good Health and Wellbeing

REFERENCES

1. Bag, R., Ghosh, M., Biswas, B., & Chatterjee, M. (2020). Understanding the spatial-temporal pattern of COVID-19 outbreak in India using GIS and India's response in managing the pandemic. (2020). *Regional Science Policy & Practice, 1063-1103.*

2. Hazbavi, Z., Mostfazadeh, R., Alaei, N., & Azizi, E. (2021). Spatial and temporal analysis of the COVID-19 incidence pattern in Iran. (2021). *Environmental Science and Pollution Research, 13605-13615.*

3. Huang, R., Liu, M., & Ding, Y. (2020). Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis. (2020). *The Journal of Infection in Developing Countries, 246-253.*

4. Sartorius, B., Lawson, A. B., & Pullan, R. L. (2021). Modelling and predicting the spatial-temporal spread of COVID-19, associated deaths and impact of key risk factors in England. (2021). *Scientific reports, 1-11*

5. Mehmood K, Bao Y, Abrar MM, Petropoulos GP, Saifullah, Soban A, Saud S, Khan ZA, Khan SM, Fahad S. "Spatiotemporal variability of COVID-19 pandemic in relation to air pollution, climate and socioeconomic factors in Pakistan". *Chemosphere. 2021 May;271:129584.*

6. Ali M, Khan DM, Aamir M, Khalil U, Khan Z (2020) Forecasting COVID-19 in Pakistan. *PLOS ONE 15(11): e0242762.*

7. Waqas, Muhammad, et al. "Analysis and prediction of COVID-19 pandemic in Pakistan using time-dependent SIR model." *arXiv preprint arXiv:2005.02353 (2020).*

8. Ahmad, I., & Muhammad Asad, S. (2020). Predictions of coronavirus COVID-19 distinct cases in Pakistan through an artificial neural network. *Epidemiology and Infection, 148, E222. doi:10.1017/S0950268820002174*

9. Shwartz-Ziv, Ravid, Itamar Ben Ari, and Amitai Armon. "Spatial-Temporal Convolutional Network for Spread Prediction of COVID-19." *arXiv preprint arXiv:2101.05304 (2020).*

References for Datasets

1. Centre for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_daily_reports]

2. Institute for Health Metrics and Evaluation (IHME) [https://covid19.healthdata.org/]

3. Timeanddate.com (T&D)

PROJECT KEY MILESTONE

Elapsed time in (days or weeks or month or quarter)since the start of the project	Mileston e	Deliverabl e
Month 1	Data Collection	Dataset with additional features
Month 2	Data Collection	Dataset with additional features
Month 3	Data Analysis	Exploratory Data Analysis & Data Visualization
Month 4	Data Analysis	Exploratory Data Analysis & Data Visualization
Month 5	SIR, Statistical Analysis and Base Paper Implementation	Analysis, Comparison of Approach, Model
Month 6	Design Model for prediction	The architecture of our model
Month 7	Model Implementation & Testing	Determine Accuracy(\Precision Accuracy Model)

Month 8	Research Paper	Ready to publish

PROJECT EQUIPMENT DETAILS

Item(s) Name	Type	No. of Units	Per Unit Cost	Total
		Total		