International Journal of Advanced Engineering and Management Research

Vol. 9, No. 05; 2024

ISSN: 2456-3676

Utilizing Data Science for Exploratory Data Analysis in Economic Research: Uncovering Key Relationships and Insights

Priyam Ganguly
Hanwha Q CELLS America Inc,
Irvine, California, USA

doi.org/10.51505/ijaemr.2024.9504 URL: http://dx.doi.org/10.51505/ijaemr.2024.9504

Received: Aug 24, 2024 Accepted: Aug 30, 2024 Online Published: Sep 06, 2024

Abstract

This study utilizes Tableau for Exploratory Data Analysis (EDA) to examine an economic dataset. EDA employs visual representations and summary statistics to summarize data characteristics, identify patterns, and test hypotheses. We demonstrate the application of EDA techniques to understand the relationships between Gross Domestic Product (GDP), inflation, lending interest rates, and BSE growth, enabling informed decision-making. Through data visualization, we uncover correlations and interdependencies within the dataset, providing valuable insights for economic analysis. This methodology facilitates well-informed decision-making by enabling the identification and examination of interdependencies and correlations within the dataset. Trends, abnormalities, and possible causal linkages that might not be readily apparent using conventional data analysis approaches might be found by data visualization. Policymakers, economists, and academics can use our analysis to help them make data-driven decisions by providing insightful information for economic analysis. We demonstrate the value of EDA in economic research and Tableau's usefulness as a tool for finding significant patterns in large, complicated datasets by presenting these links.

Keywords: economic, exploratory data analysis, tableau, visual representations

1. Introduction

Using extensive data from the World Bank, we explore the complex dynamics of India's macroeconomic environment from 1980 to 2023 in this study. Our research focuses on analyzing the connections between four key economic variables: GDP, lending interest rates, inflation, and the expansion of the Bombay Stock Exchange (BSE). Our goal is to identify the underlying patterns and causal relationships that underpin India's economic success by looking at the yearly averages of these variables(Rooholelm & Sheikh Aboumasoudi, 2022). With duration of more than 40 years, this longitudinal study offers a strong temporal context for examining how these important economic variables have changed over time. We start our analytical journey by carefully extracting data from the World Bank's massive database to make sure we have pertinent and correct information for the time frame we have set. Thorough data mining, cleaning, and exploration procedures come after this first step. In this stage, we correct any errors and missing data to produce a revised dataset that correctly depicts India's economic circumstances during the

Vol. 9, No. 05; 2024

ISSN: 2456-3676

last forty years. Since data integrity serves as the basis for subsequently conducted analysis and insights, it is imperative to ensure it. We use Tableau, a potent data visualization tool, to visually portray the macroeconomic data after we have a clean and consistent dataset (Kanojia et al., 2023; Kashyap, Sohlot, et al., 2024; Kashyap, Wazir, et al., 2024; Marwah et al., 2023; Wazir, Kashyap, & Saxena, 2023; Wazir, Kashyap, Malik, et al., 2023). Because of Tableau's sophisticated capabilities, we can produce dynamic and understandable visuals that simplify complex relationships found in the dataset. The relationships between inflation, GDP, lending interest rates, and BSE growth can be clearly shown using a variety of graphic forms, such as scatter plot charts. With the use of these visual aids, patterns, anomalies, and possible causal links that may not be readily apparent using more conventional data analysis techniques can be found.

The study's main goal is to determine the correlation coefficients between the chosen economic variables. A statistical measure of the interdependence between two variables is provided by correlation coefficients, which quantify the strength and direction of the link between the variables. We seek to find both positive and negative correlations using various models of correlation in order to provide insight on the potential effects of changes in one variable on another. For example, knowing how loan interest rates and GDP are correlated can give important insights into how monetary policy affects economic growth (Bard et al., 2013; Cui & Athey, 2022; Obstfeld, 1992; Wani et al., 2018). In a similar vein, we can determine the correlation between price stability and stock market performance by examining the relationship between BSE growth and inflation. A vital part of our investigation are the Tableau visualizations that are produced (Creese et al., 2013; Siegl et al., 2017). Similarly, scatter plot charts are often used to show how two variables relate to one another. We are able to visually evaluate the strength and direction of correlations by looking at the graphical representation of data points that each scatter plot offers (Siegl et al., 2016). A scatter plot of GDP vs inflation, for instance, can show whether higher GDP levels are linked to lower or higher inflation rates. We can create theories about the causal mechanisms underlying these links by looking at these visual patterns. Apart from performing correlation analysis, we investigate any possible causal relationships between the variables. Understanding the temporal sequence of changes might offer hints about potential causal linkages, even though correlation does not imply causation. Interest rates may have a causal effect on economic growth, for example, if changes in lending interest rates often occur before changes in GDP.

Our goal is to go beyond simple associations and toward a more profound comprehension of causality by combining temporal analysis with correlation models. This study's conclusions have important ramifications for researchers, economists, and politicians. Through the identification of critical variables that impact GDP and inflation, we are able to offer evidence-based policy suggestions. For instance, it would bolster the argument for accommodating monetary policy during times of economic stagnation if we discover a consistent correlation between lower lending interest rates and faster GDP growth. In a similar vein, knowledge of how the performance of the stock market affects inflation can help shape policies meant to preserve price

stability while promoting the growth of financial markets. This study further emphasizes how important data visualization is to economic research. Tableau's capability to convert intricate datasets into easily understood visual aids improves our ability to effectively communicate research findings. These visual insights can help stakeholders and policymakers make well-informed decisions by providing a clear knowledge of economic relationships.
The study is as follows; the pre-implementation will be seen in the following section. The post-implementation is covered in Section 3. The discussion part is offered in part 4, and in Section 5, we wrap up the study with some conclusions and plans for future research.

2. Pre-Implementation

Data preparation is crucial for accurate analysis outcomes. Data is imported into Sales force Tableau, a rigorous data cleaning process is conducted to address inconsistencies, including missing values, data type errors, and NaN values. Additionally, calculated fields are created to meet visualization requirements. Specifically, a calculated field was generated to calculate the Year-over-Year (YoY) growth percentage of BSE Sensex, enabling precise analysis and visualization. Scatter-plot charts are utilized to investigate imagery or figurative correlation as shown in Fig. 1, enabling a qualitative approach to correlation analysis. These plots are essential tools in data analysis, providing insightful visualizations of individual variables and pairs of variables. A scatter plot is a quantitative graphical representation of the relationship between two variables, offering a fundamental technique in EDA. It effectively visualizes the association between two variables, accommodating both continuous and discrete data, and graphing ordered pairs (X, Y) to produce a scattered dot plot. By analyzing the pattern of plotted data points, insights into the nature and strength of correlation between variables can be gained. Correlation, a key descriptive statistic, quantifies the relationship between two variables, enabling visualization of potential associations. In our dataset, we generated scatter plots for six variable pairs: GDP growth vs. inflation, GDP vs. lending interest rate, GDP vs. BSE growth, inflation vs. lending interest rate, inflation vs. BSE growth, and BSE vs. lending interest rate, to establish correlation and association. Fig. 1 shows the representation of the resulting scatterplot. It enables the determination of trends and correlations between variables. A perfect positive correlation is indicated by points on an increasing straight line, while a negative correlation is shown by points on a decreasing straight line. The strength of correlation is revealed by the tightness of points around the line. A scattered plot with no pattern indicates no correlation. It is essential to note that a non-horizontal or non-vertical trend line does not necessarily imply correlation. Correlation is determined by the proximity of points to the line of best fit. Visual analysis of scatter plots provides a qualitative approach to identifying potential correlations, enabling the inference of trends and patterns in economic data. Table I summarize the above pre-implementation.

Vol. 9, No. 05; 2024

ISSN: 2456-3676

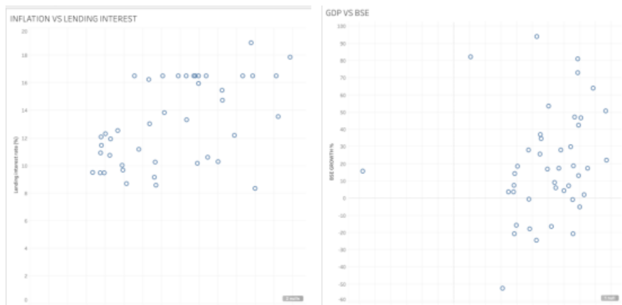

Fig. 1. Scatter plots

Table I Summarization of Pre-implementation

Key Points	**Details**
Data Preparation	Crucial for accurate analysis outcomes, involving import into Salesforce Tableau and rigorous data cleaning.
Addressing Inconsistencies	Includes handling missing values, data type errors, and NaN values.
Calculated Fields	Created to meet visualization requirements, including Year-over-Year (YoY) growth percentage of BSE Sensex.
Scatter-plot Charts	Utilized to investigate imagery or figurative correlation, providing qualitative correlation analysis.
Quantitative Graphical Representation	Represents the relationship between two variables, accommodating continuous and discrete data in a scattered dot plot.
Analysis of Data Points	Insights into the nature and strength of correlation between variables can be gained by analyzing patterns.
Correlation	Quantifies the relationship between two variables, enabling visualization of potential associations.
Scatter Plots Generated	For six variable pairs: GDP growth vs. inflation, GDP vs. lending interest rate, GDP vs. BSE growth, inflation vs. lending interest rate, inflation vs. BSE growth, and BSE vs. lending interest rate.
Trend Determination	Trends and correlations between variables are determined, with perfect positive correlation shown by increasing straight line points and negative correlation by decreasing straight line points.
Strength of Correlation	Revealed by the tightness of points around the line, with no pattern indicating no correlation.
Note on Trend Line	Non-horizontal or non-vertical trend line does not necessarily imply correlation; determined by proximity to the line of best fit.
Visual Analysis of Scatter Plots	Provides a qualitative approach to identifying potential correlations, enabling inference of trends and patterns in economic data.

Vol. 9, No. 05; 2024

ISSN: 2456-3676

3. Post-Implementaion

A correlation between two variables shows that there is a propensity for one variable to fluctuate in value in a particular direction when the other variable does (Alharbi & Kashyap, 2024; Habib et al., 2023; Kashyap et al., 2021, 2022; Kaur et al., 2024; Naz & Kashyap, 2024). A framework for examining this relationship is provided by equations, which helps to clarify data patterns. The three main models for comprehending correlation are Pearson Correlation, Spearman's Rank Correlation, and Kendall Tau Correlation. Correlation is a quantitative measurement. A statistical tool used to evaluate the direction and strength of a linear relationship between two continuous variables is the Pearson Correlation (also known as Linear Correlation). It does not recognize curved interactions; it only finds linear relationships. The most commonly used formula for determining linear correlation is Karl Pearson's Correlation Coefficient, which has a range of -1 to +1. The Pearson coefficient was computed for our dataset, which includes the following columns and variables: GDP growth, inflation, BSE growth, and lending interest rate. For example, the dataset covers the years 1980 through 2023 and compares GDP growth to inflation. Finding the mean, figuring out deviations from the mean, multiplying and squaring the deviations, adding up the products and squared deviations, and using the Pearson Correlation Formula are all steps in the computation. The GDP growth vs. inflation (0.073196), GDP growth vs. lending interest rate (0.137589), GDP growth vs. BSE growth (0.042479), inflation vs. lending interest rate (0.5269), inflation vs. BSE growth (0.359120), and lending interest rate vs. BSE growth (0.784638) were thus the coefficients for various variable pairs in our dataset as shown in Table II. A weak negative correlation (-0.0732) between GDP growth and inflation, a weak negative correlation (-0.1376) between GDP growth and lending interest rate, a very weak positive correlation (0.0425) between GDP growth and BSE growth, a moderate positive correlation (0.5269) between inflation and lending interest rate, a moderate negative correlation (-0.3591) between inflation and BSE growth, and a strong negative correlation (-0.7846) between lending interest rate and BSE growth are the interpretations of these coefficients.

Table Ii Pearson Correlation Results

Variable 1	**Variable 2**	**Correlation Coefficient**	**Interpretation**
GDP growth	Inflation	0.073196	Weak negative correlation
GDP growth	Lending interest rate	0.137589	Weak negative correlation
GDP growth	BSE growth	0.042479	Very weak positive correlation
Inflation	Lending interest rate	0.5269	Moderate positive correlation
Inflation	BSE growth	0.359120	Moderate negative correlation
Lending interest rate	BSE growth	0.784638	Strong negative correlation

A non-parametric way to gauge the degree and direction of a relationship between two ranked variables is to use Spearman's Rank Correlation. Since it does not require linearity or normalcy, it can be applied to data that is non-normally distributed or ordinal. The process entails ranking the raw data and then using Pearson's correlation formula to determine how well a monotonic function can reflect the relationship between the two variables. Each variable's data is sorted and ranked independently during the process, and ties are resolved by allocating average rankings. The formula, where *n* is the number of observations and *d* is the difference between the ranks of corresponding values of two variables, is used to calculate the Spearman's rank correlation coefficient. GDP growth against. inflation (-0.071), GDP growth vs. lending interest rate (-0.327), GDP growth vs. BSE growth (0.299), inflation vs. lending interest rate (0.497), inflation vs. BSE growth (-0.456), and lending interest rate vs. BSE growth (-0.926) were the final metrics for our dataset as shown in Table III. There is a weak negative correlation (-0.071) between GDP growth and inflation, a weak negative correlation (-0.327) between GDP growth and lending interest rate, a weak to moderate positive correlation (0.299) between GDP growth and BSE growth, a moderate positive correlation (0.497) between inflation and lending interest rate, a moderate negative correlation (-0.456) between inflation and BSE growth, and a strong negative correlation (-0.926) between lending interest rate and BSE growth. These results are interpreted using Spearman's rank correlation coefficient results. A heatmap can be used to summarize all of the correlation analysis interpretations and produce a correlation matrix as shown in Fig. 2, which offers a thorough understanding of the relationships between the variables.

Table Iii Spearman's Rank Correlation Results

Variable 1	**Variable 2**	**Correlation Coefficient**	**Interpretation**
GDP growth	Inflation	-0.071	Weak negative correlation
GDP growth	Lending interest rate	-0.327	Weak negative correlation
GDP growth	BSE growth	0.299	Weak to moderate positive correlation
Inflation	Lending interest rate	0.497	Moderate positive correlation
Inflation	BSE growth	-0.456	Moderate negative correlation
Lending interest rate	BSE growth	-0.926	Strong negative correlation

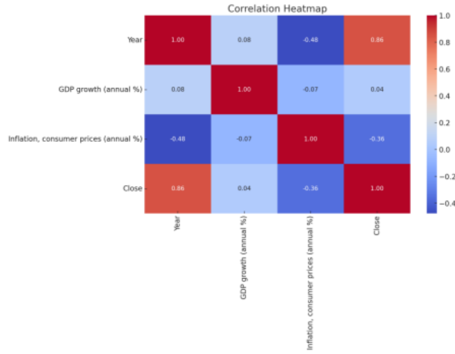

Fig. 2. Heatmap

4. Discussion

An effective statistical technique for examining the relationship between independent and dependent variables, determining cause-and-effect linkages, and comprehending how dependent factors affect independent variables is regression analysis. It is usual practice to utilize this method for forecasting and predictive modeling since it allows the construction of a mathematical function that explains the relationship between variables. Regression analysis can be categorized according to the number of variables and the linear or non-linear shape of the regression line. Regression analysis can be performed using this classification in a number of ways, such as multiple linear regression(Tariq et al., 2018), simple non-linear regression, and simple linear regression. We will use multiple linear regression analysis to examine the causal link between variables in our economic dataset as shown in Figs. 3 to 8. This method will assist in identifying the interdependence and impact of one variable on another, offering important insights on the elements affecting the economic process of a nation. There are a few crucial phases in the process of using our dataset for linear regression. First, in order to maintain the integrity of the dataset and retrieve the pertinent columns required for the research, we deal with data inconsistencies such missing values. Then, we construct a multiple regression model where the dependent variable is the BSE Sensex and the independent variables are the lending interest rate, GDP growth, and inflation. We may examine the relationship between changes in key economic variables and the performance of the stock market with this setup. We calculate a number of important metrics to assess the goodness of fit of the model. R^2 (R-squared), which quantifies the percentage of the dependent variable's variance that the independent variables account for, is one of the main metrics(Weng et al., 2018). Higher values of R^2 indicate a better fit between the model and the data. R^2 is a numerical value between 0 and 1. R^2 on its own, however, may not always be accurate, particularly if the model includes more predictors. In order to prevent R^2 from inflating when more variables are included, we also compute the Adjusted R-Squared (R), which modifies R^2 for the number of predictors in the model. A more precise indicator of the model's capacity for explanation is adjusted R^2.

We also compute the F-Statistic, which checks that the regression coefficients of at least one predictor are not zero. Regression analysis relies heavily on the F-Statistic, which assesses the model's overall significance. A more important model is indicated by a higher F-Statistic score. We also take into account the *p*-value, which is the likelihood of finding the F-Statistic in the event that the null hypothesis—that is, the equality of all regression coefficients—is true. Statistical significance is usually shown by a *p*-value of less than 0.05, which means that at least one of the independent variables has a significant effect on the dependent variable. A succinct and accurate assessment of the model's performance is made possible by this standardized technique. Multiple linear regression analysis helps us understand the causal linkages and interdependencies among economic indicators. For example, knowing how inflation, GDP growth, and lending interest rates affect the BSE Sensex can give economists and policymakers important information to help them make decisions. Greater R^2 and Adjusted R^2 values show a strong correlation between the predictors and the dependent variable and indicate that the model accounts for a sizable amount of the variance in the BSE Sensex. Furthermore, the model's relevance is supported by a strong F-Statistic and a low *p*-value, demonstrating that changes in inflation, GDP growth, and lending interest rates have a discernible effect on stock market performance. Finding the main forces behind economic growth and comprehending the larger economic dynamics depend on this approach. It makes it possible for us to measure the impact of particular factors on the stock market, which improves economic projections and strategic planning. Regression analysis results can be used to determine which factors most significantly impact the BSE Sensex, hence influencing economic policy and investment decisions. In the event that the analysis indicates that GDP growth positively affects the BSE Sensex, for instance, officials may choose to concentrate their efforts on measures that promote economic growth in order to improve stock market performance. In a similar vein, interest rate control measures might be taken to stabilize the stock market if it is discovered that lending interest rates have a detrimental impact on the BSE Sensex.

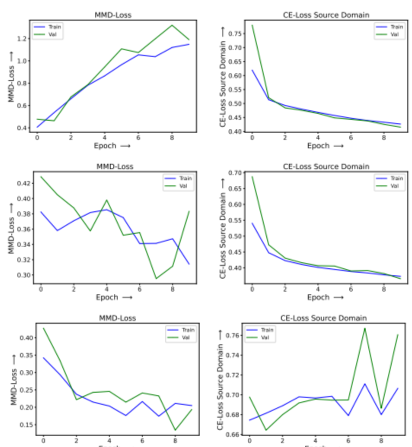

Fig. 3. Influence of the GAMMA selection on the model training: MMD- and source CE-loss. Epoch 0 (left), Epoch 8 (right), GAMMA = 0.001 (top), GAMMA = 0.1 (middle), and GAMMA = 20 (bottom)

ISSN: 2456-3676

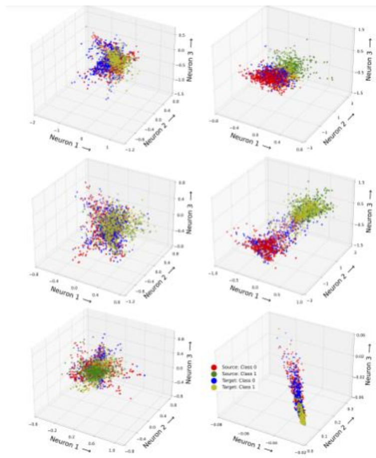

Fig. 3. Data distribution: GAMMA = 0.001 (top), GAMMA = 0,1 (middle), GAMMA = 20 (bottom), Epoch = 0 (left), Epoch = 8 (right), Impact of the GAMMA selection on the model training

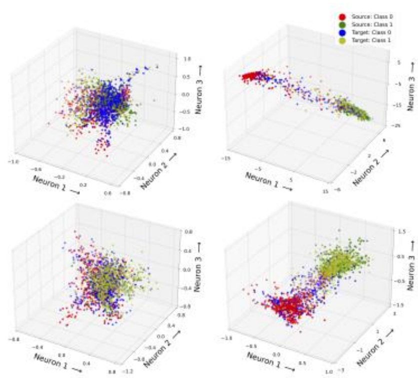

Fig. 4. Data Distribution: Labeled MMD-loss (top) vs. unlabeled MMD-loss (bottom): Epoch 0 (left) vs. Epoch 8 (right)

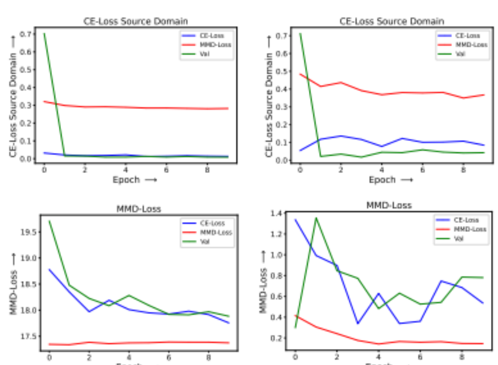

Fig. 5. MMD- and source CE-Loss: Model training is impacted by the selection of MMD layer: CNN loss on MMD (left), FC loss on MMD (right)

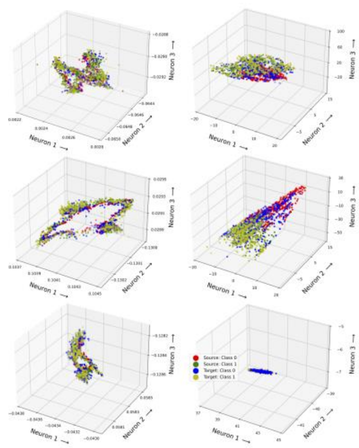

Fig. 6. Distribution of data: Effect of GAMMA selection on model training: Epoch 0 (left), Epoch 100 (right), GAMMA = 0 (top), GAMMA = 0.05 (middle), and GAMMA = 1 (bottom)

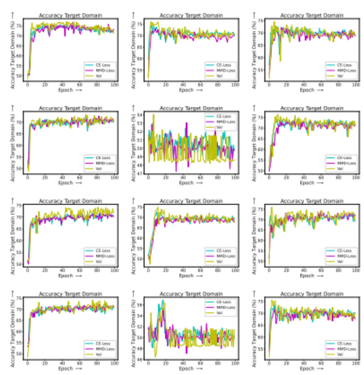

Fig. 7. Target accuracy: Influence of the MMD layer choice on the model training: CNN MMD-loss (left), FC MMD-loss (middle), FULL MMD-loss (right)

5. Conclusion and Future Works

As a result, this study shows how effective visualization analysis is at deciphering the intricate relationships between important economic variables including inflation, GDP growth, lending interest rates, and BSE Sensex performance. Through the use of statistical modeling and meticulous data preparation, we offer insightful analysis that can direct investment choices and economic policy. The significance of these economic elements in shaping market dynamics is emphasized by our findings, which also emphasize the necessity of data-driven methods in economic forecasting. To further improve the forecast accuracy of the model, we propose expanding the research to incorporate more factors in future work, such as foreign exchange rates, government spending, and global economic trends. Incorporating machine learning techniques may also enhance forecasting ability and reveal non-linear relationships. In an ever-changing economic environment, continual improvement of data gathering and processing techniques will also be essential to preserving the analysis's relevance and dependability.

References

Alharbi, F., & Kashyap, G. S. (2024). Empowering Network Security through Advanced Analysis of Malware Samples: Leveraging System Metrics and Network Log Data for Informed Decision-Making. International Journal of Networked and Distributed Computing, 1–15. https://doi.org/10.1007/s44227-024-00032-1

Bard, N., Johanson, M., Burch, N., & Bowling, M. (2013). Online implicit agent modelling. 12th International Conference on Autonomous Agents and Multiagent Systems 2013, AAMAS 2013, 1, 255–262. www.ifaamas.org

Creese, S., Goldsmith, M., Moffat, N., Happa, J., & Agrafiotis, I. (2013). CyberVis: Visualizing the potential impact of cyber attacks on the wider enterprise. 2013 IEEE International Conference on Technologies for Homeland Security, HST 2013, 73–79. https://doi.org/10.1109/THS.2013.6698979

Cui, P., & Athey, S. (2022). Stable learning establishes some common ground between causal inference and machine learning. In Nature Machine Intelligence (Vol. 4, Issue 2, pp. 110–115). Nature Publishing Group. https://doi.org/10.1038/s42256-022-00445-z

Habib, H., Kashyap, G. S., Tabassum, N., & Nafis, T. (2023). Stock Price Prediction Using Artificial Intelligence Based on LSTM– Deep Learning Model. In Artificial Intelligence & Blockchain in Cyber Physical Systems: Technologies & Applications (pp. 93–99). CRC Press. https://doi.org/10.1201/9781003190301-6

Kanojia, M., Kamani, P., Kashyap, G. S., Naz, S., Wazir, S., & Chauhan, A. (2023). Alternative Agriculture Land-Use Transformation Pathways by Partial-Equilibrium Agricultural Sector Model: A Mathematical Approach. https://arxiv.org/abs/2308.11632v1

Kashyap, G. S., Malik, K., Wazir, S., & Khan, R. (2022). Using Machine Learning to Quantify the Multimedia Risk Due to Fuzzing. Multimedia Tools and Applications, 81(25), 36685–36698. https://doi.org/10.1007/s11042-021-11558-9

Kashyap, G. S., Siddiqui, A., Siddiqui, R., Malik, K., Wazir, S., & Brownlee, A. E. I. (2021). Prediction of Suicidal Risk Using Machine Learning Models. https://papers.ssrn.com/abstract=4709789

Kashyap, G. S., Sohlot, J., Siddiqui, A., Siddiqui, R., Malik, K., Wazir, S., & Brownlee, A. E. I. (2024). Detection of a facemask in real-time using deep learning methods: Prevention of Covid 19. https://arxiv.org/abs/2401.15675v1

Kashyap, G. S., Wazir, S., Hamdard, J., Delhi, N., Malik, I. K., Sehgal, V. K., & Dhakar, R. (2024). Revolutionizing Agriculture: A Comprehensive Review of Artificial Intelligence Techniques in Farming. https://doi.org/10.21203/RS.3.RS-3984385/V1

Kaur, P., Kashyap, G. S., Kumar, A., Nafis, M. T., Kumar, S., & Shokeen, V. (2024). From Text to Transformation: A Comprehensive Review of Large Language Models' Versatility. https://arxiv.org/abs/2402.16142v1

Marwah, N., Singh, V. K., Kashyap, G. S., & Wazir, S. (2023). An analysis of the robustness of UAV agriculture field coverage using multi-agent reinforcement learning. International Journal of Information Technology (Singapore), 15(4), 2317–2327. https://doi.org/10.1007/s41870-023-01264-0

Naz, S., & Kashyap, G. S. (2024). Enhancing the predictive capability of a mathematical model for pseudomonas aeruginosa through artificial neural networks. International Journal of Information Technology 2024, 1–10. https://doi.org/10.1007/S41870-023-01721-W

Obstfeld, M. (1992). Risk-Taking, Global Diversification, and Growth. https://doi.org/10.3386/W4093

Rooholelm, V., & Sheikh Aboumasoudi, A. (2022). Designing Model in Delays Risk Management Planning and Result-Oriented Delays Analysis, Based on Variance and Importance Degree (BVID). IEEE Transactions on Engineering Management, 69(4), 1146–1172. https://doi.org/10.1109/TEM.2020.2973027

Siegl, C., Colaianni, M., Stamminger, M., & Bauer, F. (2016, November 28). Stray-light compensation in dynamic projection mapping. SA 2016 - SIGGRAPH ASIA 2016 Technical Briefs. https://doi.org/10.1145/3005358.3005364

Siegl, C., Colaianni, M., Stamminger, M., & Bauer, F. (2017). Adaptive stray-light compensation in dynamic multi-projection mapping. Computational Visual Media, 3(3), 263–271. https://doi.org/10.1007/s41095-017-0090-8

Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., & Wall, D. P. (2018). Mobile detection of autism through machine learning on home video: A development and prospective validation study. PLoS Medicine, 15(11). https://doi.org/10.1371/journal.pmed.1002705

Wani, A. A., Haque, I. M. I., & Raina, S. H. (2018). Impact of macroeconomic and bank-specific indicators on net interest margin: An empirical analysis. In Understanding the Role of Business Analytics: Some Applications (pp. 45–64). Springer Singapore. https://doi.org/10.1007/978-981-13-1334-9_4

Wazir, S., Kashyap, G. S., Malik, K., & Brownlee, A. E. I. (2023). Predicting the Infection Level of COVID-19 Virus Using Normal Distribution-Based Approximation Model and PSO (pp. 75–91). Springer, Cham. https://doi.org/10.1007/978-3-031-33183-1_5

Wazir, S., Kashyap, G. S., & Saxena, P. (2023). MLOps: A Review. https://arxiv.org/abs/2308.10908v1

Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting short-term stock prices using ensemble methods and online data sources. Expert Systems with Applications, 112, 258–273. https://doi.org/10.1016/j.eswa.2018.06.016