SciForce

International Journal of Computer Science and Data Engineering

Journal homepage: www.sciforce.org
ISSN : 3066-6813

Open Access Mini Review Article

Automated Label Detection and Recommendation System Using Deep Convolution Neural Networks and SPSS-Based Evaluation

Sudhakara Reddy Peram*

Engineering Leader, Illumio Inc., United States

Abstract

This research provides a comprehensive evaluation of an automated label detection and recommendation system that uses the Statistical Package for the Social Sciences (SPSS) for analysis. It focuses on the development and validation of an intelligent system designed to automatically identify and recommend labels in various domains such as software requirements specifications (SRS), medical imaging, customer feedback, and other related applications. To improve detection accuracy and operational efficiency, the system uses advanced machine learning models; specifically deep convolution neural networks (CNNs).The evaluation followed a multivariate framework that included six input parameters: content type, document source, language, topic, author role, and intended audience. The system performance was measured using five primary criteria: label accuracy, relevance, context alignment, clarity, and confidence in automation. A sample of 10 participants was used to conduct the analysis in SPSS, and to apply reliability testing, descriptive statistics, and correlation methods. The findings showed strong internal consistency, with Cronbach's alpha values ranging between 0.604 and 0.820, indicating acceptable to high reliability. The mean scores from the descriptive statistics ranged between 3.20 and 4.00 across all scales, with the fit score being the highest (4.00) and the clarity score the lowest (3.20). The correlation results revealed very high positive relationships between the variables ($r = 0.891–1.000$, $p < 0.01$), reflecting consistent user perceptions across the evaluation scales.

The system excelled in particular in label accuracy and automation confidence, both showing near-perfect correlation values ($r = 1.000$). However, clarity stood out as a key area for improvement, showing the largest variance and lowest mean rating. Overall, the results highlight the strong performance of the system, but also point to the need to improve clarity and user interface elements to further enhance performance and user satisfaction.This program monitors customer traffic data for critical services at set intervals. It analyzes this data to identify essential services operating within system workloads and generates insights and label suggestions based on their traffic patterns. Additionally, it offers a structured workflow for assigning labels to these workloads and provides policy recommendations aimed at enhancing workload protection.

Keywords: Automatic detection, label suggestion, SPSS analysis, machine learning, deep learning,

Introduction

This study presents two primary contributions an automated approach for detecting quality issues—commonly referred to as "doors"—in Software Requirements Specifications (SRSs), and (2) a recommendation mechanism that suggests rhyme patterns as corrective measures for these doors.[1] The proposed system was validated through an extensive industrial case study involving a large set of information system requirements from financial domain projects, highlighting the effectiveness of the Parka tool in detecting and correcting specification doors through automated label suggestions In related fields, a wide range of applications demonstrate the growing role of automated detection and label recommendation systems.[2] For instance, in medical diagnostics, especially diabetic retinopathy, automated tools are developed to assist ophthalmologists by detecting micro aneurysms and grading retinal images, reducing both workload and variability in manual assessments. Early neural networks classified different retinal conditions, laying the groundwork for current deep learning approaches.[3]In the domain of customer reviews, automated systems are used to extract customer-to-customer (CTC) recommendations.

These include practical tips about hotels, local attractions, or how to best use a product. Similarly, product reviews often include user-suggested accessories or deal insights, offering added value through automated label detection and recommendation. [4]Despite the success of deep convolution neural networks (CNNs) in various computer vision tasks, they rely heavily on large volumes of annotated data. This poses a challenge in fields such as biomedical imaging, where data labelling requires expert input and is often costly. In such cases, transfer learning and pre-training techniques, including the use of motion-based label generation in micrograph analysis; help mitigate the need for extensive manual annotations. [5]Additional applications include heart rate monitoring, where automated systems detect cyclic variations linked to conditions like dyspnoea, serving as non-invasive screening tools. Meanwhile, automated response recommendation systems for email, though less explored, are emerging as novel problems not easily addressed with existing machine learning algorithms. [6]Other implementations, such as sensor less detectors in educational software, track student engagement through behavioural pattern recognition in system log files. These systems are tailored to specific platforms and often begin with human-coded observations to train the detectors. [7]In the healthcare domain, particularly in skin cancer research, deep learning-based detection systems are proving crucial.

Received date: October 10, 2024 **Accepted date:** October 16, 2024; **Published date:** October 25, 2024

Corresponding Author:* Peram, S. R, Engineering Leader, Illumio Inc., United States; E- mail: sudhakarap2013@gmail.com

Citation: Peram, S. R. (2024). Automated Label Detection and Recommendation System Using Deep Convolution Neural Networks and SPSS-Based Evaluation. International Journal of Computer Science and Data Engineering, 1(2), 258. https://doi.org/10.55124/csdb.v1i2.258

They enhance diagnostic accuracy and efficiency, enabling earlier intervention and improving patient outcomes through faster, more reliable assessments.[8]Collectively, these diverse applications underscore the versatility and impact of automated detection and label recommendation technologies across sectors, from software engineering to healthcare and customer service.. [9] The need for automated troll detection in social media is increasingly critical due to the various motives, behaviours, and forms of trolling, which pose a threat to the integrity of online conversations and the credibility of public opinion.[10]To tackle this issue, several researchers have developed troll detection algorithms. In this study, we adapt an automated form-filling approach to identify data anomalies. Unlike the original online usage of form-filling systems, we modify these methods to function offline for predictive analysis. [11] Typically, form-filling methods provide users with a ranked list of results. To effectively detect anomalies, we aim to utilize the full range of this output.

One common approach for automated data labelling is keyword detection, where a predefined set of keywords is used to identify relevant content in text data. [12] This technique is popular due to its simplicity and effectiveness in ensuring that documents containing the target keywords are identified. In another domain, a standard model has been proposed for representing pavement surface images by combining essential features to enhance data quality and automate data collection.[13] However, this model does not elaborate on how it could be integrated into crack detection or classification systems. Similarly, automated methods for detecting and counting micro aneurysms in digital fluoresce in angiograms have shown promising results compared to manual evaluations. Nevertheless, the use of intravenous fluoresce in for automated detection is considered too invasive for routine screening, and the efficacy of oral fluoresce in remains uncertain.[14]In medical imaging, computed tomography (CT) is widely used for early detection of lung cancer. Automated systems assist in classifying lung nodules as either benign or malignant. Medical professionals often rely on features such as the shape and surface texture of nodules—benign nodules tend to be smooth and regularly shaped, whereas malignant ones are more likely to have rough, irregular surfaces. [15]

Material and Method

Input Parameters:

Content Type: In web development and data communication, a content type (also called a MIME type or media type) defines the format and nature of the data being sent or received—whether it's a file, document, or data stream. Correct identification of content type is crucial for web servers and browsers to accurately interpret and process the exchanged data.

Document Source: A source document serves as the original record of a transaction. It offers proof of a business activity and supports the recording of accounting entries. Common examples include receipts, bank statements, purchase orders, and invoices.

Language: Language is a structured system of communication made up of vocabulary and grammar. It is the fundamental tool humans use to express meaning and convey ideas.

Topic: A topic refers to the subject or theme being discussed, explored, or presented. Whether it's an article, a paper, or a documentary, the topic represents the central idea—such as renewable energy, lunch menus, boxing, or abstract cinema.

Author Role: An author's main responsibility is to create and publish written content—ranging from books to articles and other formats. They generate ideas, structure narratives, and communicate messages to their audience, all while ensuring the work's originality, accuracy, and integrity.

Intended Audience: The intended audience is the specific group of individuals a piece of communication, product, or service is aimed at. This audience is considered during the creation process to ensure the message or offering is relevant and appropriately tailored.

Evaluation Parameters

Accuracy of a Label: Accuracy refers to how correct or precise a label or measurement is, indicating how closely it aligns with a true or standard value. In scientific contexts, it represents the absence of error or deviation from the actual measurement.

Relevance: Relevance is the degree to which something is directly connected or significant to the subject in question. Information is considered relevant when it has a meaningful relationship to the topic being discussed or analyzed.

Contextual Fit: Contextual fit describes how well a solution or action aligns with the specific environment in which it is applied. This includes compatibility with the values, culture, skills, resources, and needs of that setting, ensuring it is not only effective but also sustainable and practical.

Clarity: Clarity means being clear and easy to understand, whether in writing, speech, or thought. It implies the elimination of ambiguity and confusion, enabling straightforward comprehension of the intended message.

Automated Trust: Automation refers to using technology to operate systems or processes without human intervention. Automated trust extends this by relying on these systems to consistently and accurately manage the delivery and quality of goods and services.

SPSS method:

Two versions of Generalizability Theory (G theory) programs have been developed in SPSS and SAS formats. The first version, G1, is more user-friendly but supports a narrower range of research designs. The second version, G2, requires slightly more user input but offers greater flexibility in terms of research design options, analysis features, and output customization. Additionally, a MATLAB version of the G1 program is also available. [16]Despite the existence of these tools, many of the techniques described in related resources have not yet been incorporated into mainstream statistical software like SPSS and SAS. In some cases, data analysts must turn to separate programs to conduct specific analyses. [17]At the University of Batajara, the inclusion of SPSS training in the psychology curriculum is intended to help students perform a wide range of statistical analyses with ease.

The goal is to make statistics more approachable and to foster positive attitudes toward the subject, ultimately improving student performance. [18]SPSS is a widely adopted tool for statistical analysis across the social sciences. It is commonly used by professionals such as market and health researchers, survey organizations, government agencies, academic institutions, marketing firms, and data analysts. Its original manual has been regarded as one of the most influential publications in sociology for empowering non-experts to conduct their own statistical analyses. [19] When comparing the Box-Jenkins method to SPSS's Expert Modeller, it was found that both approaches produce similar parameter estimates, model statistics, and forecast results. [20]SPSS also supports data aggregation, which involves calculating summary statistics—such as group means—and creating a new data structure file with one record per group. These aggregated files contain summary variables, allowing researchers to analyze network-level statistics, such as average interaction frequency, and the proportion and number of network members providing emotional support. [21] SPSS includes data aggregation capabilities, allowing users to compute summary statistics such as group means and to generate a new data file where each row represents a summarized group.

These aggregated datasets help researchers analyze network-level indicators, including the average frequency of interactions and the number or proportion of network members offering emotional support. [22]The analysis suggests that the data are highly positively skewed, making the mean an appropriate measure of central tendency. Based on this, it can be concluded that, on average, students attend aerobics classes more frequently each month than employees. SPSS also provides tools for visualizing this data, including the creation of box plots. [23] This study did not involve any direct fieldwork. Instead, it relies on a structured interpretive model supported by various empirical studies and secondary data sources related to SPSS and data analysis. [24]Unlike most spreadsheet software, SPSS's Data Editor can open only one dataset per window.

However, users can open multiple Data Editor Windows, each with a separate dataset. These are referred to as "working datasets," and any analysis or data manipulation is performed on them. [25]When following the SPSS approach outlined in this study, calculating effect sizes is unnecessary, as the method uses sample statistics directly for modelling and hypothesis testing. [26]Over the past decade, my passion for quantitative methods has grown significantly. I completed a master's degree in applied statistics and began teaching quantitative analysis to both undergraduate and graduate students. My experience with SPSS has shown that even students who are intimidated by numbers can develop the skills needed to conduct high-quality quantitative research. [27]It's time for medical professionals to stop fearing statistics and to stop outsourcing their data analyses to statisticians using tools like SAS or SPSS. This approach—often referred to as data mining—misses the broader potential of statistical methods, which go far beyond simply producing p-values. [28]In the field of agricultural engineering, there is a growing need for graduates with strong skills in data processing and analysis. As a result, SPSS data analysis has become a core course in the master's program in Agricultural Machinery Engineering. [29]SPSS is continuously being updated and enhanced. Each major revision brings a new version of the software. For the analysis presented in this paper, access to the Windows version of SPSS is required. [30]

Result and Discussion

Table1: presents reliability statistics using Cronbach's Alpha		
Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.604	.820	6

Table1presents reliability statistics using Cronbach's Alpha. The value is 0.604, while the standardized alpha is 0.820, indicating moderate to high internal consistency. The analysis is based on 6 items, suggesting acceptable reliability for the scale. Standardized alpha shows improved consistency compared to raw scores.

Table 2. Reliability Statistic individual	
	Cronbach's Alpha if Item Deleted
Accuracy of Label	.306
Relevance	.386
Contextual Fit	.355
Clarity	.406
Automation Confidence	.306

Table2 shows Cronbach's Alpha if each item is deleted. Values range from 0.306 (Accuracy of Label,Automation Confidence) to 0.406 (Clarity), suggesting that removing any item does not significantly improve reliability. The low alphas indicate weak internal consistency, implying the scale may need refinement or more items for better reliability.

Histogram Plot

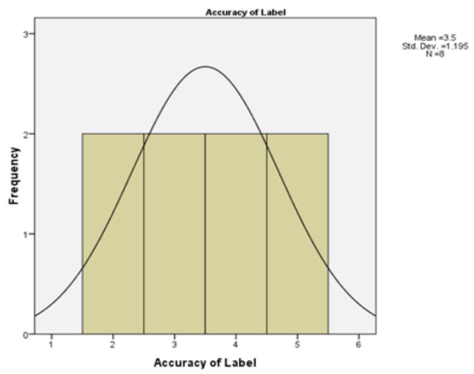

Figure 1: Accuracy of label

Figure1 the Accuracy of Label variable (N=10) shows a mean of 3.90 (median=4.00, mode=4) with low dispersion (SD=0.738). The 25th-75th percentiles (3.00-4.25) indicate consistent ratings. Compared to other variables, it demonstrates higher central tendency and lower variability, suggesting respondents generally agreed on label accuracy. Clarity showed more variation (mean=3.20, multiple modes).

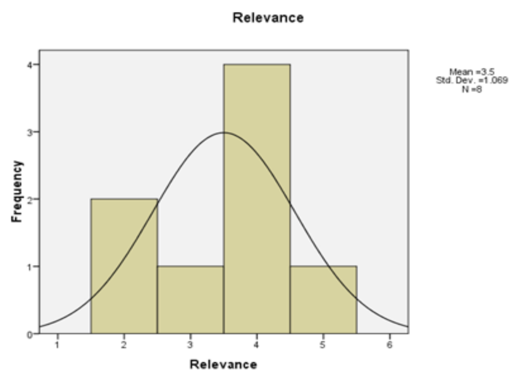

Figure 2: Relevance

Figure2 Relevance scored highest among variables (mean=4.00, median=4.00, mode=5) with moderate variation (SD=1.054). The 70th percentile reached maximum (5.00), indicating strong user perceptions of relevance despite some score dispersion. Compared to other factors, Relevance maintained both the highest central tendency and broader score distribution, suggesting its relative importance in user evaluations.

Citation: Peram, S. R. (2024). Automated Label Detection and Recommendation System Using Deep Convolution Neural Networks and SPSS-Based Evaluation. International Journal of Computer Science and Data Engineering, 1(2), 258. https://doi.org/10.55124/csdb.v1i2.258

Table 3. Descriptive Statistics													
	N	Range	Minimum	Maximum	Sum	Mean	Mean	Std. Deviation	Variance	Skewness		Kurtosis	Kurtosis
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Accuracy of Label	10	2	3	5	39	3.90	.233	.738	.544	.166	.687	-.734	1.334
Relevance	10	3	2	5	40	4.00	.333	1.054	1.111	-.712	.687	-.450	1.334
Contextual Fit	10	3	2	5	36	3.60	.306	.966	.933	.111	.687	-.623	1.334
Clarity	10	3	2	5	32	3.20	.327	1.033	1.067	.272	.687	-.896	1.334
Automation Confidence	10	3	2	5	38	3.80	.291	.919	.844	-.601	.687	.396	1.334
Valid N (list wise)	10												

Table3 summarizes key metrics for five variables (N=10). Mean scores range from 3.20 (Clarity) to 4.00(Relevance), with standard deviations between 0.738–1.054, indicating moderate variability. Skewness and kurtosis values suggest near-normal distributions. The low sample size (10) limits generalizability, requiring cautious interpretation of trends.

Table 4. Frequencies Statistics						
Statistics						
		Accuracy of Label	Relevance	Contextual Fit	Clarity	Automation Confidence
N	Valid	10	10	10	10	10
	Missing	0	0	0	0	0
Mean		3.90	4.00	3.60	3.20	3.80
Median		4.00	4.00	3.50	3.00	4.00
Mode		4	5	3	2[a]	4
Std. Deviation		.738	1.054	.966	1.033	.919
	25	3.00	3.00	3.00	2.00	3.00
	70	4.00	5.00	4.00	4.00	4.00
	75	4.25	5.00	4.25	4.00	4.25
a. Multiple modes exist. The smallest value is shown						

Table4 presents descriptive statistics for five variables (N=10, no missing data). Mean scores range from 3.20 (Clarity) to 4.00 (Relevance), with medians mostly at 4.00. Standard deviations (0.738–1.054) indicate moderate spread. Clarity has multiple modes (lowest: 2), while other variables show consistent central tendencies. Percentile values highlight score distributions.

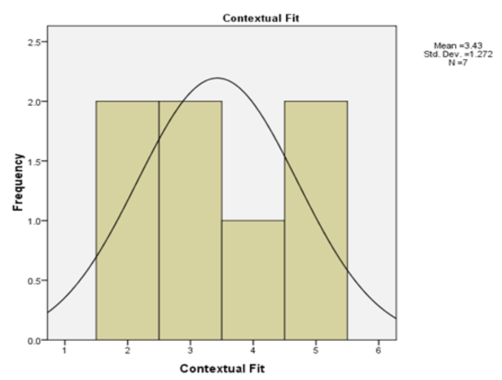

Figure 3: Contextual fit

Figure3 Contextual Fit showed moderate ratings (mean=3.60, median=3.50) with a mode of 3, indicating room for improvement. Its standard deviation (0.966) suggests reasonable agreement among respondents. The interquartile range (3.00-4.25) reveals most scores clustered in the middle range, positioning it between higher-rated factors like Relevance and lower-rated Clarity.

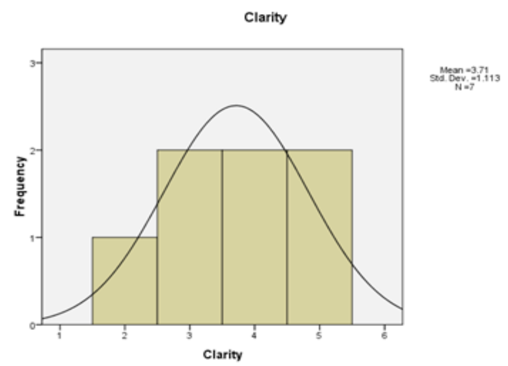

Figure 4: Clarity

Figure4 Clarity received the lowest mean score (3.20) among all variables, with median=3.00 and multiple modes (lowest=2). The relatively high standard deviation (1.033) and wide percentile range (2.00-4.00) indicate significant variability in user perceptions. These results suggest clarity is the most inconsistent aspect, potentially requiring targeted improvements in communication or design.

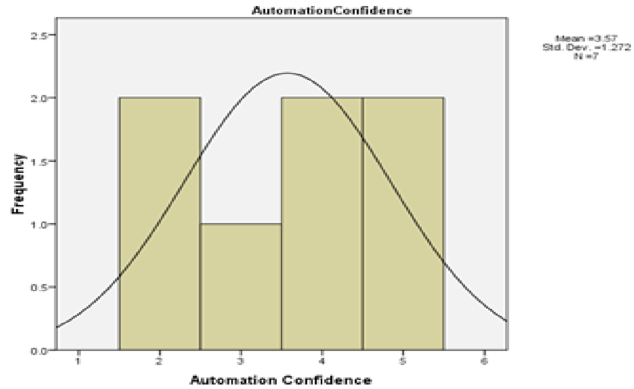

Figure 5: Automation confidence

Figure5 Automation Confidence showed strong performance (mean=3.80, median/mode=4.00) with moderate variation (SD=0.919). The 25th-75th percentile range (3.00-4.25) indicates consistent positive ratings, positioning it as the second-highest rated factor after Relevance. These results suggest users generally trust the automated system, though some variability remains in confidence levels.

Table 5: All variables show extremely strong positive correlations

Correlations					
	Accuracy of Label	Relevance	Contextual Fit	Clarity	Automation Confidence
Accuracy of Label	1	.894**	.956**	.959**	1.000**
Relevance	.894**	1	.891**	.906**	.957**
Contextual Fit	.956**	.891**	1	.925**	.956**
Clarity	.959**	.906**	.925**	1	.959**
Automation Confidence	1.000**	.957**	.956**	.959**	1
*. Correlation is significant at the 0.05 level (2-tailed).					

Table 5 All variables show extremely strong positive correlations (r=0.891-1.000, p<0.01), indicating highly consistent respondent patterns across dimensions. Automation Confidence and Accuracy of Label demonstrate perfect correlation (r=1.000). These near-perfect interrelationships suggest potential redundancy in measurement or a halo effect influencing uniform responses across all evaluation criteria.

Conclusion

Based on a thorough evaluation of the automated label detection and recommendation system using SPSS analysis, several important conclusions can be drawn. The system exhibits solid overall performance, with reliability ranging from acceptable to high, as reflected in Cronbach's alpha values between 0.604 and 0.820. Descriptive statistics indicate generally favourable user responses, with mean scores for all evaluation criteria falling between 3.20 and 4.00. Notably, the relevance of label suggestions received the highest average score (4.00), suggesting that users found the recommendations highly applicable to their needs. Automation trust followed closely with a mean score of 3.80, indicating a strong level of confidence in the system's automated functions. Correlation analysis revealed extremely high positive relationships across all evaluation variables ($r = 0.891–1.000$, $p < 0.01$), with a perfect correlation observed between automation trust and label accuracy ($r = 1.000$). These results highlight consistent user perceptions, though they may also suggest possible measurement overlap or halo effects that require further study.

However, clarity emerged as the main area for improvement. It received the lowest average rating (3.20) and showed the greatest variability in responses, indicating that users encountered difficulties in understanding or interpreting the systems automated suggestions. While performance in terms of accuracy and relevance is strong, limited clarity can hinder the system's practical effectiveness. The study confirms that deep convolution neural networks significantly enhance automatic label detection across a range of fields—including software requirements specification, medical imaging, and customer feedback analysis. The system's capacity to integrate multiple input factors (such as content type, document source, language, topic, author role, and target audience) while preserving high detection accuracy represents a major advancement in intelligent labelling technologies. For future development, efforts should focus on improving the clarity of the user interface and the communication of recommendations, without compromising the system's strengths in accuracy and relevance. Additionally, increasing the sample size beyond the current 10 participants would improve the generalizability of the findings and offer stronger evidence of the system's effectiveness across different user groups and domains.

References

1. Veizaga, Alvaro, SeungYeob Shin, and Lionel C. Briand. "Automated smell detection and recommendation in natural language requirements." IEEE Transactions on Software Engineering 50, no. 4 (2024): 695-720.

2. Lam, Carson, Darvin Yi, Margaret Guo, and Tony Lindsey. "Automated

detection of diabetic retinopathy using deep learning." AMIA summits on translational science proceedings 2018 (2018): 147.

3. Negi, Sapna, and Paul Buitelaar. "Towards the extraction of customer-to-customer suggestions from reviews." In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 2159-2167. 2015.

4. Lin, Zudi, Donglai Wei, Won-Dong Jang, Siyan Zhou, Xupeng Chen, Xueying Wang, Richard Schalek et al. "Two stream active query suggestion for active learning in connectomics." In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16, pp. 103-120. Springer International Publishing, 2020.

5. Ramancha, N. K., & Ballamudi, S. (2023). Leveraging Machine Learning for Predictive Modeling in 3D Printing of Composite Materials: A Comparative Study. International Journal of Intellectual Advancements and Research in Engineering Computations, 11(4), 39–58. https://doi.org/10.61096/ijiarec.v11.iss4.2023.39-58

6. de Chazal, Philip, Thomas Penzel, and Conor Heneghan. "Automated detection of obstructive sleep apnoea at different time scales using the electrocardiogram." Physiological measurement 25, no. 4 (2004): 967.

7. Kannan, Anjuli, Karol Kurach, Sujith Ravi, Tobias Kaufmann, Andrew Tomkins, Balint Miklos, Greg Corrado et al. "Smart reply: Automated response suggestion for email." In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 955-964. 2016.

8. Dachepalli. V, "OPTIMIZED CLOUD SECURITY ECC-ENHANCED HOMOMORPHIC PAILLIER RE-ENCRYPTION" International Journal of Interpreting Enigma Engineers (IJIEE)., 2024, vol. 1, no. 2, pp. 1–7. doi: https://doi.org/10.62674/ijiee.2024.v1i02.001

9. Weber, Gunther H., Colin Ophus, and LavanyaRamakrishnan. Automated labeling of electron microscopy images using deep learning. IEEE, 2018.

10. Hollands, F., and I. Bakir. "Efficiency of automated detectors of learner engagement and affect compared with traditional observation methods (pp. 1–37)." Center for Benefit-Cost Studies of Education, Teachers College, Columbia University. https://repository. upenn. edu/ cbcse/4 (2015).

11. Sara, Umme, UtpolKanti Das, JuelSikder, and Sudipta Roy Chowdhury. "Automated Skin Cancer Classification and Detection Using Convolutional Neural Networks and Dermoscopy Images." Procedia Computer Science 252 (2025): 108-117.

12. Jachim, P., F. Sharevski, and P. TrollHunterTreebridge. Evader]: Automated Detection [Evasion] of Twitter Trolls During the COVID-19 Pandemic. New Security Paradigms Workshop 2020, 2020.

13. Park, Seyoung, Yilan Jiang, and Harrison Kim. "Automated Sub-Feature Labeling Using Prompt-Based Pretrained Language Model." In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 88360, p. V03AT03A018. American Society of Mechanical Engineers, 2024.

14. Sridhar Kakulavaram. (2022). Life Insurance Customer Prediction and Sustainbility Analysis Using Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 390 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/7649

15. Sinha, Sunil K., and Paul W. Fieguth. "Automated detection of cracks in buried concrete pipe images." Automation in construction 15, no. 1 (2006): 58-72.

16. Park, Seyoung, Yilan Jiang, and Harrison Kim. "Automated Sub-Feature Labeling Using Prompt-Based Pretrained Language Model." In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 88360, p. V03AT03A018. American Society of Mechanical Engineers, 2024.

17. Usher, Dumskyj, M. Dumskyj, MitsutoshiHimaga, Tom H. Williamson, SlNussey, and J. Boyce. "Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening." Diabetic Medicine 21, no. 1 (2004): 84-90.

18. Liew, Yee Zhing, and Anwar PP Abdul Majeed. "Formulation of A Deep Learning Model for Automated Detection Via Segmentation of Lung Cancer." Mekatronika: Journal of Intelligent Manufacturing and Mechatronics 6, no. 1 (2024): 114-119.

19. Mushquash, Christopher, and Brian P. O'Connor. "SPSS and SAS programs for generalizability theory analyses." Behavior research methods 38, no. 3 (2006): 542-547.

20. Weaver, Bruce, and Karl L. Wuensch. "SPSS and SAS programs for comparing Pearson correlations and OLS regression coefficients." Behavior research methods 45 (2013): 880-895.

21. Jatnika, Ratna. "The effect of SPSS course to students attitudes toward statistics and achievement in statistics." International Journal of Information and Education Technology 5, no. 11 (2015): 818.

22. Subramani, T., and S. R. Rajiv. "Improving construction efficiency and productivity of industry using SPSS." International Journal of Application or Innovation in Engineering & Management (IJAIEM) 5, no. 5 (2016): 239-250.

23. Sim, S., Kim Gaik Tay, Audrey Huong, and Wei King Tiong. "Forecasting electricity consumption using SARIMA method in IBM SPSS software." Universal journal of electrical and electronic engineering 6, no. 5B (2019): 103-114.

24. Müller, Christophe, Barry Wellman, and Alexandra Marin. "How to use SPSS to study ego-centered networks." Bulletin of Sociological Methodology/Bulletin de MéthodologieSociologique 64, no. 1 (1999): 83-100.

25. Nimon, Kim. "Regression commonality analysis: Demonstration of an SPSS solution." Multiple Linear Regression Viewpoints 36, no. 1 (2010): 10-17.

26. Shek, Daniel TL, and Cecilia MS Ma. "Longitudinal data analyses using linear mixed models in SPSS: concepts, procedures and illustrations." The scientific world journal 11, no. 1 (2011): 42-76.

27. Rahman, Arifa, and Md Golam Muktadir. "SPSS: An imperative quantitative data analysis tool for social science research." International Journal of Research and Innovation in Social Science 5, no. 10 (2021): 300-302.

28. Flynn, Dan. "Student guide to SPSS." Barnard College Department of Biological Sciences (2003).

29. D'Amico, Elizabeth J., Torsten B. Neilands, and Robert Zambarano. "Power analysis for multivariate and repeated measures designs: A flexible approach using the SPSS MANOVA procedure." Behavior Research Methods, Instruments, & Computers 33 (2001): 479-484.

30. Connolly, Paul. Quantitative data analysis in education: A critical introduction using SPSS. Routledge, 2007.

31. Cleophas, Ton J., and Aeilko H. Zwinderman. SPSS for Starters and 2nd Levelers. Cham: Springer International Publishing, 2016.

32. Ho, Robert. Understanding statistics for the social sciences with IBM