Annexure

REFERENCES

A. Journals and Books

- Syed and Kumar (2021). Elmaghraby, A. IoT in Smart Cities: "A Survey of Technologies, Practices and Challenges. Smart Cities", 4, 429–475. https://doi.org/ 10.3390/smartcities4020024

- Moazzami, Majid, Shahvand, Niloufar, Kabalci, Ersan, Shahinzadeh, Kabalci, Yasin, Gharehpetian and Gevork. (2021). "Internet of Things Architecture for Intelligent Transportation Systems in a Smart City". 285-290. 10.1109/GPECOM52585.2021.9587692.

- Lakshmi Shankar Iyer (2021). "AI enabled applications towards intelligent transportation, Transportation Engineering", Volume 5,100083, ISSN 2666-691X, https: //doi.org /10.1016 / j. treng.2021.100083.

- Agarwal, Gurjar, Birla "Application of Artificial intelligence for development of intelligent transport system in smart cities" Int. J. Transp. Eng. Traffic Syst., 1 (2015), pp. 20-30.

- Abbas, S., et al. (2011). "Bio-inspired neuro-fuzzy based dynamic route selection to avoid traffic congestion". International Journal of Scientific and Engineering Research, 2(6), pp. 284-289.

- Abigail R., Jorge B., Franklin L., Angel S and Santiago M. (2019) "Integrated Information System for Urban Public Transport" International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958, Volume-8 Issue-4C, April 2019 pp. 97 – 102.

- Abduljabbar, R., Dia, H., Liyanage, S., and Bagloee, S. A. (2019). "Applications of artificial intelligence in transport: An overview". Sustainability, pp. 11, 189.

- Al-Dweik, A., Muresan, R., Mayhew, M., and Lieberman, M. (2017). "IoT-based multifunctional Scalable real-time Enhanced Road Side Unit for Intelligent Transportation Systems". In Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on, pp. 1-6.

- Albino, V., Berardi, U., and Dangelico, R. M. (2015). "Smart cities: Definitions, dimensions, performance, and initiatives". Journal of Urban Technology, 22(1), pp. 3-21.

- Ammara, U., Rasheed, K., Mansoor, A., Al-Fuqaha, A., and Qadir, J. (2022). "Understanding the Dynamics and Outcomes of Smart Cities: The Role of Interdependencies and Stakeholder Interactions". In Proceedings of the International Conference on Smart City Applications, pp. 1-10.

- Amadeo, M., Campolo, C., Quevedo, J., Corujo, D., Molinaro, A., Iera, A., ... & Vasilakos, A. V. (2016). "Information-centric networking for the Internet of Things: Challenges and opportunities". IEEE Network, 30(2) pp. 92-100.

- Bai, M., Lin, Y., Ma, M., Wang, P., and Duan, L. (2020). PrePCT: "Traffic congestion prediction in smart cities with relative position congestion tensor. Neurocomputing". https://doi.org/10.1016/j.neucom.2020.08.075 pp. 146-159.

- Balasubramanian, S. B., Balaji, P., Munshi, A., Almukadi, W., Prabhu, T. N., K, V., & Abouhawwash, M. (2023). "Machine learning based IoT system for secure traffic management and accident detection in smart cities". PeerJ Computer Science, 9, e1259. https://doi.org/10.7717/peerj-cs.1259, pp.39-56.

- Bhardwaj, S., Malik, H., Chauhan, D. S., & Rana, R. (2021). "A comprehensive survey on IoT-based traffic prediction models for smart cities". Journal of Ambient Intelligence and Humanized Computing, 12(10), pp. 14873-14892.

- Cai, C. (2009). "Adaptive Traffic Signal Control Using Approximate Dynamic Programming (Ph.D. Thesis)". University College, London, UK.

- Chowdhury, A. (2016). "Priority-based and secured traffic management system for an emergency vehicle using IoT". In Engineering & MIS (ICE), International Conference on pp. 1-6.

- Cunha, M., Amaral, A., Silva, A., and Pinheiro, C. (2021). "Smart public transportation for smart cities: A systematic literature review. Sustainability", 13(10), 5441. doi:10.3390/su13105441. pp 109-117.

- Collotta, M., Sun, Y., Di Persio, L., Ebeid, E. S. M., & Muradore, R. (2018). "Smart Green Applications: From Renewable Energy Management to Intelligent Transportation Systems. Energies", pp.11, 1317.

- Carmona, M. (2010). "Public places, urban spaces: The dimensions of urban design. Routledge".

- Chakraborty, A. D. A., Kumar, A., Roy, A., Roy, S., Chakraborty, D., Saha, H. N., et al. (2019). "Intelligent traffic control system: Towards smart city". IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), https://doi.org/10.1109/IEMCON.2019.8936188 pp. 1123-1129.

- Chong, H. F., and Ng, D. W. K. (2016). "Development of IoT device for traffic management system". In Research and Development (SCOReD), 2016 IEEE Student Conference on pp. 1-6.

- Coutard, O., et al. (2014). "Urban Megatrends: Towards a European research agenda" pp.1-17.

- Dresner, P. S. K. (2008). "A Multiagent approach to Autonomous Intersection Management". Journal of Artificial Intelligence Research, 31, pp. 591–656.
- Dou E. (2015), "AI and ML Based Public transport Information center" University of Wollongong Thesis Collection Chapter 2.3 to 2.6 pp 12 – 30.
- Deak, G., & Walravens, N. (2019). "The future of urban mobility: Towards integrated smart mobility solutions. Transport Reviews", 39(2), pp 157-181.
- Dimitrakopoulos, G., & Bravos, G. (2016). "Current Technologies in Vehicular Communication". Springer. pp.39-52.
- Doolan, R., & Muntean, G.-M. (2016). "EcoTrec—A novel VANET-based approach to reducing vehicle emissions". IEEE Transactions on Intelligent Transportation Systems, 18(3), pp.608–620.
- Eken S. and Sayar A. (2014) "A smart bus tracking system based on location-aware services and QR codes," 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, pp. 299-303
- Erlmann, S., & Dantzig, S. V. (2018). Exploring the potential of shared mobility services to alleviate urban traffic congestion: A literature review. Transport Reviews, 38(6), pp. 769-793.
- Ersoy, P., and Börühan, G. (2015). "Intelligent Transportation Systems and Their Applications in Road Transportation Industry in Turkey". In Proceedings of the 12th International Conference on Logistics & Sustainable Transport, Celje, Slovenia, pp. 11–13 June 2015.
- Ferrara, A., Sacone, S., and Siri, S. (2018). "Fundamental of traffic dynamics. In Freeway traffic modelling and control" (Advances in Industrial Control Series, Vol. 1185, pp. 44).
- Fintikakis, N., & Bourka, A. (2018). "Commuting patterns and smart cities: A comparative study of Stockholm and London". TeMA Journal of Land Use, Mobility and Environment, 11(3), pp. 239-254.
- Gkiotsalitis, K., & Cats, O. (2018). "Assessing the impact of autonomous vehicles on our cities. Transportation Research Part A: Policy and Practice", pp. 118, 407-417.
- Gohar, M., Sagheer, A., Javaid, N., Anpalagan, A., & Khan, F. (2019). "Fog computing-enabled intelligent transportation system: Architecture, opportunities, and challenges". IEEE Access, 7, pp. 29037-29057.
- Giffinger, R., and Gudrun, H. (2010). "Smart cities ranking: An effective instrument for the positioning of cities? ACE Architecture, City and Environment", 4(12), pp. 7–25.

- Ghasem-Aghaee, N., Tuncer Oren, and Levent Yilmaz. (2019). "Simulation, Intelligence and Agent: Exploring the Synergy. In Current and Future Developments in Artificial Intelligence". Bussum, The Netherlands: Bentham Science Books Publisher. pp.1-58.

- Goggin, G. (2012). "Driving the internet: Mobile internets, cars, and the social. Future Internet", 4(1), pp. 306-321.

- Hall, R. E., and Pfeiffer, U. (2016). "The impact of smart cities on travel behavior: An exploratory study. Transportation Research Part A: Policy and Practice", 88, pp. 96-114.

- Han, Q., Liang, S., & Zhang, H. (2015). "Mobile cloud sensing, big data, and 5G networks make an intelligent and smart world". IEEE Network, 29(2), pp. 40-45.

- Iker, Z., Alessandro, S., & Saioa, A. (2016). "Smart city concept: What it is and what it should be. Journal of Urban Planning and Development", 142(1), 04015005. pp. 178-189.

- Iyer L. (2021) "AI enabled applications towards intelligent transportation" Transportation Engineering 5 (2021) 100083 open access article under the CC BY-NC-ND license Published by Elsevier Ltd. pp 1 – 11.

- Jacobsen, R. H., Aliu, D., & Ebeid, E. (2017). "A Low-Cost Vehicle Tracking Platform Using Secure SMS". In Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security, Porto, Portugal, pp. 24–26.

- Jacobsen, R. H., Gabioud, D., Basso, G., Alet, P. J., Azar, A. G., & Ebeid, E. (2015). "SEMIAH An Aggregator Framework for European Demand Response Programs". In Proceedings of the Euromicro Conference on Digital System Design (DSD), Madeira, Portugal, 26–28 August 2015; pp. 470–477.

- Jarrah M. and Shrida F. (2017) "A multi-objective evolutionary solution to improve the quality of life in smart cities," 2017 14th International Conference on Smart Cities: Improving Quality of Life Using ICT and IOT (HONET-ICT), pp 36 - 39.

- Jung, C., and Couclelis, H. (2017). "Smart cities and the politics of urban data". Journal of Planning Education and Research, 37(4), pp. 457-466.

- Janahan, S. K., Veeramanickam, M., Arun, S., Kumar, N., Anandan, R., & Javed, S. (2018). "IoT-based smart traffic signal monitoring system using vehicle counts". International Journal of Engineering and Technology, https://doi.org/10.14419/ijet.v7i2.21.12388 pp. 7,309.

- Jin, J., Ma, X., & Kosonen, I. (2017). "An intelligent control system for traffic lights with simulation-based evaluation. Control Engineering Practice," https://doi.org/10.1016/j.conengprac.2016.09.009 pp. 58, 24-33.

- Joo, H., Ahmed, S. H., & Lim, Y. (2020). "Traffic signal control for smart cities using reinforcement learning. Computer Communications". https://doi.org/10.1016/j.comcom.2020.03.005 pp. 111 - 124

- Kaur, A., and Kaur, P. (2017). "Internet of Things for transportation: A review". IEEE Sensors Journal, 17(14), pp.4371-4378.

- Khan, A., Salah, K., and Zeadally, S. (2019). "Fog computing for intelligent transportation systems: Opportunities and challenges. Computer Networks," 160, pp. 96-112.

- Khallouk, A., Echab, H., Ez-Zahraouy, H., and Lakouari, N. (2018). "Traffic flow behaviour at Un-signalized intersection with crossing pedestrians. Physics Letters A," 382, pp. 566–573.

- Kikuchi, S. (2009). "Artificial intelligence in transportation analysis: Approaches, methods, and applications". Transportation Research Part C, 17(5), pp. 455.

- Kuberkar and Singhal (2020) "Factors Influencing Adoption Intention of AI Powered Chatbot for Public Transport Services within a Smart City" International Journal on Emerging Technologies 11(3): 948-958(2020) ISSN No. (Print): 0975-8364 ISSN No. (Online): 2249-3255 pp. 948 – 958.

- Kumar, N., Rahman, S. S., & Dhakad, N. (2020). "Fuzzy inference enabled deep reinforcement learning-based traffic light control for intelligent transportation systems". IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2020.2984033 pp. 147-169

- Kitchin, R. (2014). "Big Data, new epistemologies and paradigm shifts." Big Data & Society, 1(1), 2053951714528481. pp. 45-63.

- Lakshmi, I. (2016). "A literature survey on Big Data Analytics in Service Industry". International Journal of Engineering and Computer Science, 5(4).

- Lee, R., Blogg, M., Myers, E., Kyte, M., Dixon, M., List, G., Flannery, A., Troutbrck, R., Brilon, W., Werner, B., et al. (2007). "National Cooperative Highway Research Program (NCHRP)" Report 572 (Vol. 9). Washington, DC, USA: Transportation Research Board of the National Academics. pp. 327-346.

- Lee, W.-H., & Chiu, C.-Y. (2020). "Design and implementation of a smart traffic signal control system for smart city applications". Sensors, 20(2), https://doi.org/10.3390/s20020508 pp. 508

- Liu, J., Li, J., Niu, X., Cui, X., & Sun, Y. (2014). GreenOCR: "An Energy-Efficient Optimal Clustering Routing Protocol". Computer Journal, 58, pp. 1344–1359.

- Lingani, G. M., Rawat, D. B., & Garuba, M. (2019). "Smart traffic management system using deep learning for smart city applications". IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), https://doi.org/10.1109/CCWC.2019.8666539 pp. 0101–0106.

- Mamoona Humayun, Sadia Afsar, Maram Fahaad Almufareh, N. Z. Jhanjhi, Mashayel AlSuwailem, "Smart Traffic Management System for Metropolitan Cities of Kingdom Using Cutting Edge Technologies", Journal of Advanced Transportation, vol. 2022, Article ID 4687319,13. https:// doi.org/10.1155/ 2022/ 4687319. pp. 197-213.

- Marko Pavelić, Zvonimir Lončarić, Marin Vuković, Mario Kušek. "Internet of Things Cyber Security: Smart Door Lock System". 978-1-5386-7189-4/18. pp. 213-222.

- Misbahuddin, S., Zubairi, J. A., Saggaf, A., Basuni, J., Sulaiman, A., and Al-Sofi, A. (2015). "IoT Based Dynamic Road Traffic Management for Smart Cities". In Proceedings of the 2015 12th International Conference on High-Capacity Optical Networks and Enabling/Emerging Technologies (HONET), Islamabad, Pakistan, 21–23 December 2015; pp. 142–146.

- Mukti, I. Y., & Prambudia, Y. (2018). "Challenges in Governing the Digital Transportation Ecosystem in Jakarta: A Research Direction in Smart City Frameworks". Challenges, pp. 9, 14.

- M. S. Kamal, S. Parvin, K. Saleem, H. Al-Hamadi, and A. Ghawanmeh. "Efficient low cost supervisory system for internet of things enabled smart home". In2017 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 864–869.

- Misuraca, G., Mureddu, F., & Osimo, D. (2014). "Policy-making 2.0: Unleashing the power of big data for public governance". Open Government, Springer, pp. 171-188.

- Mondal, M. A., & Rehena, Z. (2019). "Intelligent traffic congestion classification system using artificial neural network". Companion Proceedings of the World Wide Web Conference, https://doi.org/10.1145/3308560.3317053 pp. 110–116.

- Malek, G., Li, C., Yang, Z., Hasan, N. A. H., & Zhang, X. (2012). "Improved the Energy of Ad hoc On-Demand Distance Vector Routing Protocol". In International Conference on Future Computer Supported Education, Published by Elsevier, IERI, pp.355-361.

- Neelakandan, S. (2016). "Large scale optimization to minimize network traffic using MapReduce in big data applications". In International Conference on Computation of Power, Energy Information and Communication (ICCPEIC) https://doi.org/10.1109/ICCPEIC.2016.7557196 pp.193-199.

- Neelakandan, S., & Paulraj, D. (2020). "An Automated Exploring And Learning Model For Data Prediction Using Balanced CA-Svm". Journal of Ambient Intelligence and Human Computing, 1-12. Springer, Berlin pp.1868–5137.

- Nasr, E., Kfoury, E., and Khoury, D. (2016). "An IoT approach to vehicle accident detection, reporting, and navigation". In Multidisciplinary Conference on Engineering Technology (IMCET), IEEE International, pp. 231-236.

- Neha Patil, Shrikant Ambatkar and Sandeep Kakde. "IoT Based Smart Surveillance Security System using Raspberry Pi" International Conference on Communication and Signal Processing, April pp. 6-8.

- Navarathna, Pramathi J., and Vindhya P. Malagi. "Artificial intelligence in smart city analysis." 2018 International conference on smart systems and inventive technology (ICSSIT). IEEE, 2018. pp. 96-104.

- Ou, H., Zhang, J., & Wang, Y. (2016). "Development of intelligent traffic control system based on Internet of Things and FPGA technology in Proteus. Traffic", pp. 20-22.

- Pattanaik, V., Singh, M., Gupta, P. K., & Singh, S. K. (2016). "Smart real-time traffic congestion estimation and clustering technique for urban vehicular roads". In IEEE Region 10 Conference (TENCON) , pp. 3420-3423.

- Prakash N. and Anudeep P. (2018), "Intelligent Passenger Information System Using IoT for Smart Cities" AISC, volume851 pp. 67 – 76.

- Pradhan M. and Kumar U. (2019), "Machine Learning using Python", WILEY publication First Edition, pp. 179 – 181.

- Pereira, F., Rodrigues, J. J., and Saleem, K. (2018). "IoT-based transportation systems: A survey on applications, challenges, and opportunities". Journal of Network and Computer Applications, 110, pp. 40-54.

- Pau, G., Campisi, T., Canale, A., Severino, A., Collotta, M., and Tesoriere, G. (2018). "Smart Pedestrian Crossing Management at Traffic Light Junctions through a Fuzzy-Based Approach". Future Internet, pp.10, 15.

- Psomakelis, E., et al. (2012). "Big IoT and Social Networking Data for Smart Cities" pp. 176 - 193

- Rahman A. and Ajala (2018) "Conceptualising Smart City for the Development of Nigeria's Urban Transportation" Advances In Multidisciplinary and Scientific Research Vol. 4 No. 2, June 2018 pp. 65 – 72.

- Raj G Anvekar and Rajeshwari M Banakar. "IoT application development: Home security system". 2017 IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2017) pp.159-172.

- Ravi Kishore Kodali, Vishal Jain, Suvadeep Bose, Lakshmi Boppana."IoT based smart security and home automation system". International Conference on Computing, Communication and Automation (ICCCA2016) pp. 188-192.

- Rizwan, P., Suresh, K., and Rajasekhara Babu, M. (2016). "Real-time smart traffic management system for smart cities by using the Internet of Things and big data". In Emerging Technological Trends (ICETT), International Conference on, pp. 1-7.

- Ramachandra, S. H., Reddy, K. N., Vellore, V. R., Karanth, S., & Kamath, T. (2016). "A novel dynamic traffic management system using on-board diagnostics and Zigbee protocol". In Communication and Electronics Systems (ICCES), International Conference on pp. 1-6.

- Rego, A., Garcia, L., Sendra, S., & Lloret, J. (2018). "Software-defined Network-based control system for efficient traffic management for emergency situations in smart cities". Future Generation Computer Systems, pp. 88, 243-253.

- Satpathy, S., Mohan, P., Das, S., et al. (2020). "A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA". Journal of Supercomputing, https://doi.org/10.1007/s11227-019-03013-2 pp. 76, 5849-5861.

- Saeed, Y., et al. (2016). "Impact of Cognition on User Authentication Scheme in Vehicle using Fuzzy Logic and Artificial Neural Network". International Journal of Computer Science and Information Security, pp.14(10), 285.

- Sundar G. (2017), "IoT based Passenger Information System Optimized for Indian Metros" International Conference on Electronics, Communication and Aerospace Technology ICECA 2017 pp. 92 – 96.

- Sarrab, M., Pulparambil, S., & Awadalla, M. (2020). "Development of an IoT based real-time traffic monitoring system for city governance". Global Transitions, 2 https://doi.org/10.1016/j.glt.2020.09.004. pp. 230-245.

- Silva, F. M., & Silva, R. F. (2019). "Congestion charging and smart cities: A systematic review". Transportation Research Part A: Policy and Practice, 121, pp. 331-350.

- Sochor, J., & Dvorak, J. (2020). "Impact of smart city solutions on urban mobility: A systematic literature review. Sustainability," 12(7), 2771 pp. 314-323.

- Subramani, Neelakandan & Berlin, M. & Tripathi, Sandesh & Devi, V. & Bhardwaj, Indu & Natarajan, Arulkumar. (2021). "IoT-based traffic prediction and traffic signal control system for smart city. Soft Computing." 25. 10.1007/s00500-021-05896-x. pp. 152-160.

- Sharma, V., You, I., Pau, G., Collotta, M., Lim, J. D., & Kim, J. N. (2018). "LoRaWAN-Based Energy-Efficient Surveillance by Drones for Intelligent Transportation Systems". Energies, pp.11, 573.

- Sun, Y., and Song, H. (2017). "Secure and Trustworthy Transportation Cyber-Physical Systems". Springer. pp.103-121.

- Subbulakshmi, P. and Prakash, M. (2018). "Mitigating eavesdropping by using fuzzy-based MDPOP-Q learning approach and multilevel Stackelberg game theoretic approach in wireless CRN". Cognitive Systems Research, pp. 52, 853-861.

- Tarawneh, M., Alzyoud, F., & Sharrab, Y. (2023). "Artificial Intelligence Traffic Analysis Framework for Smart Cities".

- Tchuitcheu, W. C., Bobda, C., & Pantho, M. J. H. (2020). "Internet of smart cameras for traffic lights optimization in smart cities. Internet of Things". https://doi.org/10.1016/j.iot.2020.100207 pp. 32-47.

- Thakur, T. T., Naik, A., Vasari, S., & Gogate, M. (2016). "Real-time traffic management using the Internet of Things". In Communication and Signal Processing (ICCSP), 2016 International Conference on pp. 1950-1953.

- Theodoridis, E., Mylonas, G., & Chatzigiannakis, I. (2013). "Developing an IoT smart city framework". In Information, Intelligence, Systems and Applications (Lisa), 2013 Fourth International Conference on pp. 1-6.

- Troutbeck, R. (2016). "Modelling Signalized and Un-Signalized Junctions". In Handbook of Transport Modelling (pp. 443–460). Bingley, UK: Emerald Group Publishing Limited.

- Vakula D. and Raviteja B. (2017), "Smart Public Transport for Smart Cities" Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017) IEEE Xplore Compliant - Part Number:CFP17M19-ART, ISBN:978-1-5386-1959-9 pp. 805 – 810.

- Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2014). "Short-term traffic forecasting: Where we are and where we're going". Transportation Research Part C: Emerging Technologies, 43, pp. 3-19.

- Wen, W. (2008). "A dynamic and automatic traffic light control expert system for solving the road congestion problem". Expert Systems with Applications, 34, pp. 2370–2381.

- World Health Organization. (2018). "Global Status Report on Road Safety 2018". Geneva, Switzerland: World Health Organization.

- Yao, H., Gao, P., Wang, J., Zhang, P., Jiang, C., & Han, Z. (2019). "Capsule network assisted IoT traffic classification mechanism for smart cities". IEEE Internet of Things Journal, 6(5), pp. 7515-7525.

- Zafar, N., Ul Haq, I., Sohail, H., Chughtai, J.-U.-R., & Muneeb, M. (2022). "Traffic prediction in smart cities based on hybrid feature space". IEEE Access, 10, https://doi.org/10.1109/ACCESS.2022.3231448. pp.134333-134348.

- Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). "Internet of Things for smart cities". IEEE Internet of Things Journal, 1(1), pp. 22-32.

- Zhang, J., Wang, X., & He, J. (2020). "Fog computing-based intelligent transportation systems: Challenges and solutions". IEEE Transactions on Intelligent Transportation Systems, 21(7), pp. 2909-2924.

- Zhou, H., Wang, X., Ma, J., Zhang, J., & Ma, L. (2020). "An IoT-based traffic prediction model for smart cities using deep learning algorithms. Sensors," 20(6), 1763. pp. 142-149.

- Kumar, Ratan. (2023). "A research paper on AI for traffic management". 10.13140/RG.2.2.21119.69287.

- Sreelatha, Mahalakshmi and Yadav(2023). "Artificial Intelligence Based Autonomous Traffic Regulator", Proceedings of CEUR Workshop Proceedings, New Delhi, India.

B. Websites

[1] https://www.slideshare.net/paojean2000/urban-transport-problems Accessed on 19 - 08 - 2021

[2] https://rideamigos.com/smart-mobility-in-smart-cities/ Accessed on 19 - 08 – 2021

[3] https://seoulsolution.kr/en/content/7664 /Accessed on 19 - 08 - 2021

[4] http://ethesis.usm.my/jspui/handle/123456789/12348 Accessed on 19 - 08 – 2021

[5] https://www.digiteum.com/iot-data-collection/ Accessed on 19.08.2021

[6] https://www.intelligenttransport.com/transportarticles/21458/city-public-transportation-india/ Accessed on 19.08.2021

[7] https://www.economicsdiscussion.net/sales/sales-forecasting-methods/32270 Accessed on 15.09.2021

[8] https://iq.opengenus.org/advantages-and-disadvantages-of-linear-regression/ Accessed on 15.09.2021

[9] https://towardsdatascience.com/prediction- engineering-how-to-set-up-your-machine- learning-problem-b3b8f622683b/ Accessed on 15.09.2021

[10] https://www.marketresearch.com/Schonfeld-Associates-Inc-v417 Ratios- Budgets-13373044/ Accessed on 20.09.2021

[11] https://www.keboola.com/blog/linear- regression-machine-learning / Accessed on 20.09.2021

[12] https://internal.ncl.ac.uk/ask/numeracy-maths-statistics/statistics/regression-and- correlation/coefficient-of-determination-squared.html / Accessed on 27.09.2021

[13] http://core.ecu.edu/psyc/wuenschk/docs30/EffectSizeConventions.pdf / Accessed on 27.09.2021

[14] https://www.obviously.ai/post/the-difference-between-training-data-vs-test-data-in- machine-learning / Accessed on 25.02.2022

[15] https://data.cityofchicago.org/ Accessed on 20.03.2023

[16] https://chicagotraffictracker.com/ Accessed on 10.09.2023.

[17] https://data.cityofchicago.org/Transportation/Chicago-Traffic-Tracker-Historical Congestion-Esti/kf7e-cur8/about_data/ Accessed on 10.09.2023.

[18] https://www.tomtom.com/products/route-monitoring/ Accessed on 01.10.2023.

[19] https://www.tomtom.com/products/traffic-stats/ Accessed on 01.10.2023.

[20] https://www.tomtom.com/products/junction-analytics/ Accessed on 12.10.2023.

[21] https://en.wikipedia.org/wiki/Feature_selection / Accessed on 20.02.2023

[22] https://towardsdatascience.com/feature-selection-and-dimensionality-reduction-f488d1a035de / Accessed on 25.05.2023

[23] https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62 /Accessed on 10.06.2023.

[24] https://c3.ai/glossary/machine- learning/precision/#:~:text= Precision%20is %20one%20indicator%20of, the%20number%20of%20false %20positives / Accessed on 12.06.2023

[25] https://www.shiksha.com/online-courses/articles/sensitivity-vs-specificity/ #:~: text= Sensitivity%20measures%20how%20well%20a, have%20 significantly%20fewer%20False%20Negatives. /Access on 12.06.2023

[26] https://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7255#:~: text=Definition,positives%20which%20are%20correctly %20identified / Accessed on 28.07.2023

[27] https://developers.google.com/machine-learning/crash-course/ classification /roc-and-auc /Accessed on 30.07.2023

[28] https://www.koshegio.com/chi-square-online-calculator / Accessed on 18.12.2023

[29] https://www.koshegio.com/t-test-distribution-free-calculator/Accessed on 20.12.2023

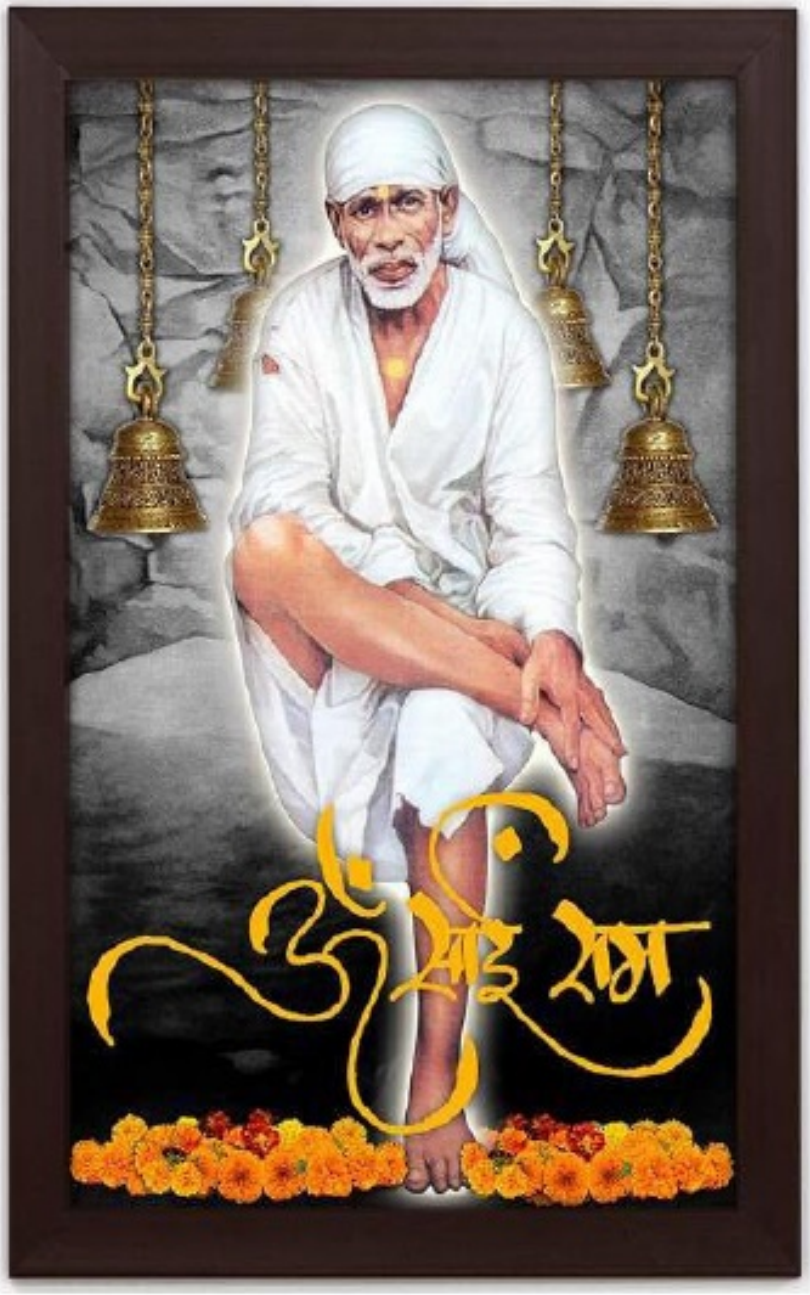
[30] https://www.koshegio.com/f-distribution-calculator-graph/Accessed on 20.12.2023

DISSEMINATION OF RESEARCH WORK

[P1] Avinash Dangwani, Dr. Ashok kumar Jetawat and Dr. Jayshree Jain (2023), " Exploring the Role of Machine Learning Algorithms for Smart Commuting in Smart Cities", 38th Indian Engineering Congress Conference of The Institution of Engineering(IEI) India, held at Jabalpur 27-29 December 2023.

[P2] Avinash Dangwani, Dr. Ashok kumar Jetawat and Dr. Chandansingh Rawat (2023), "Use of AI in Cloud based Certificate Authentication for Travel Concession", 4th International Springer Conference on Mobile Computing and Sustainable informatics(ICMSCI 2023) Organized by Tribhuvan University Nepal, 11 – 12 January 2023. Research paper published in Lecture Notes on Data Engineering and Communications Technologies Springer Nature Singapore in May 2023 ISSN 2367-4512, 2367-4520 (electronic), ISBN 978-981-99-0834-9, 978-981-99-0835-6 (eBook). Indexed by SCOPUS, INSPEC, EI Compendex and UGC Care List Group – II.

[P3] Avinash Dangwani and Dr. Chandansingh Rawat (2021), "Data Analytics Sales Prediction Model", 12th International Conference on Sustainable Global Trends: Planet, People and Profit by Pacific Institute of Management and Pacific Business School Udaipur, 16 – 17 April 2021. Research paper published in Pacific Business Review International Journal in Aug 2021 ISSN: 0974-438X August, 2021. Indexed by Web of Science Group and UGC Care List Group – II.


ॐ साईं राम

HYPOTHESIS -1 SURVEY QUESTIONS

1. Age of Participant.

2. Gender

3. City of Residence:

4. Are you aware of the concept of smart cities ?
    a) Yes
    b) No

5. Have you used any smart transportation technologies in your city ?
    a) Yes
    b) No

6. Which of the following smart transportation technologies being used ?
    a) AI-Based Systems
    b) Fog Computing
    c) IoT-Based Traffic Prediction Models
    d) Machine Learning-Based Traffic Prediction Models

7. How would you rate the reliability of public transportation in your city ?
    a) Low
    b) Moderate
    c) High
    d) Very High

8. How would you rate the effectiveness of smart transportation technologies in enhancing the transportation system for smart cities ?
    a) Low
    b) High

9. Do you have any additional comments or suggestions regarding smart transportation technologies in smart cities ?

Published Research Papers

The Institution of Engineers (India)

Established 1920, Incorporated by Royal Charter 1935

A Century of Service to the Nation

38th Indian Engineering Congress

December 27-29, 2023

Certificate of Paper Presentation

This is to certify that

Avinash Dangwani, Ashok Kumar Jetawat, Jayshree Jain

Department of Computer, Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan

Have participated in the 38th Indian Engineering Congress held at Jabalpur during December 27-29, 2023 and presented a paper titled **'Exploring the Role of Machine Learning Algorithms for Smart Commuting in Smart Cities'.**

Maj Gen (Dr) MJS Syali, VSM (Retd)
Secretary & Director General

PRESIDENT

Technical Volume

38th Indian Engineering Congress

December 27-29, 2023

Venue: Hotel Royal Orbit, Jabalpur

Theme

Reimagining Tomorrow: Shaping the Future through Disruptive and Interdisciplinary Technologies

Organised by

The Institution of Engineers (India)

8 Gokhale Road, Kolkata

Hosted by

Jabalpur Local Centre

CV_IEC 2023 TV 011

Exploring the Role of Machine Learning Algorithms for Smart Commuting in Smart Cities

Avinash Dangwani*, Ashok Kumar Jetawat and Jayshree Jain

Department of Computer, Pacific Academy of Higher Education and Research University, Udaipur

✉ avin861@gmail.com*

Abstract: *Cities are developing with boundless boundaries, Policymakers and urban organizers are investigating ways to address smart Infrastructure in smart cities such as smart water management, Environmental monitoring, Green spaces, smart waste management, smart commuting, smart solution to traffic congestion and hassle-free day-by-day transportation. Smart commuting may be a key component of the smart framework, which points to making strides in efficiency, security, sustainability, and by and large quality of life in urban and country regions. Smart commuting leverages innovation and data-driven solutions to upgrade different angles of the transportation framework. Development in innovation such as IoT, Artificial intelligence, and Machine learning has invented various solutions for smart commuting in smart cities. Smart cities contribute capital in the present day and effective public transportation frameworks, including buses, metro, and rail systems. These frameworks regularly incorporate Traffic Monitoring, real-time tracking, ticket sales prediction, remote door unlocking, seat availability prediction and other advanced innovations to make commuting more helpful.Our paper investigate and discuss various ML algorithms and techniques to address traffic congestionproblems in smart cities using traffic prediction methods.*

Keywords: *Classifiers;tomtom; Precision; Recall; TP Rate; FP Rate; Bayes; Naïve Bayes; Random Forest*.

INTRODUCTION

Machine learning algorithms can be very effective in designing models for solving traffic congestion problems. They can help with various aspects of traffic management and optimization by analyzing historical data, making predictions, and making recommendations.

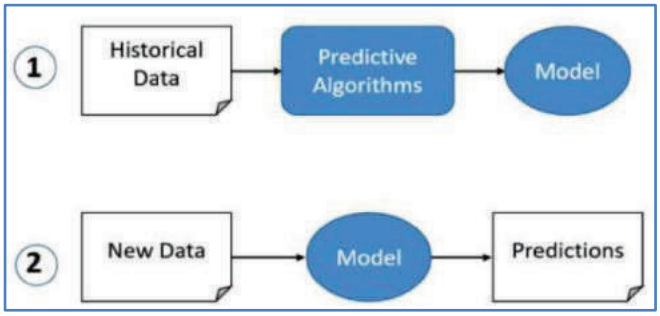


Figure 1 Model Devising

Some of the common cases are Traffic predictions, Traffic flow management, Demand Management, Parking optimization. Traffic prediction and control systems in smart cities are essential for managing urban congestion and improving overall transportation efficiency. Various machine learning and IoT-based models have been developed to address these challenges. Some of the existing approaches in this field are shown below.

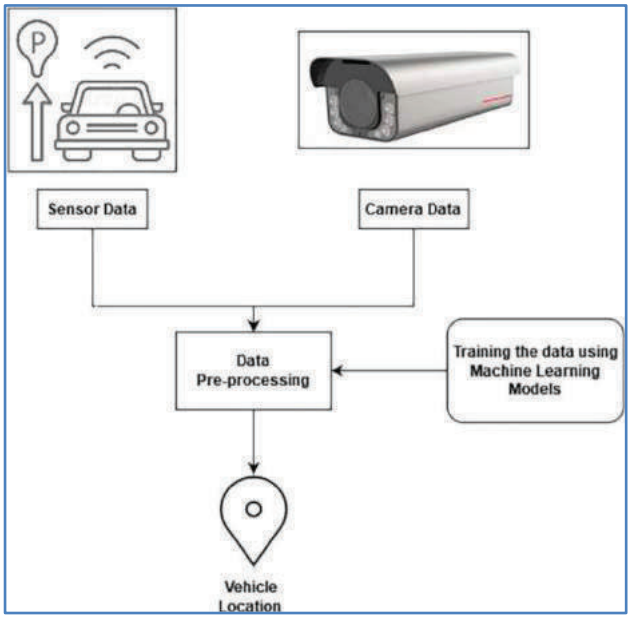


Figure 2 Data Processing Model (Courtesy : peerj.com)

IoT devices and sensors deployed across road networks can collect real-time data on traffic flow, vehicle speeds, and occupancy. By analysing this data, traffic prediction models can provide accurate and up-to-date traffic forecasts. Using real-time traffic data collected from IoT sensors, adaptive traffic signal control systems can adjust signal timings based on current traffic conditions. These systems aim to minimize congestion and improve traffic flow efficiency. By combining machine learning-based traffic prediction models with control algorithms, predictive traffic control systems can anticipate traffic conditions and adjust signal timings proactively. These systems help prevent traffic jams before they happen. Predictive machine learning models for smart transportation systems integrate data from various sources, including traffic cameras, GPS devices, and weather forecasts, to predict traffic congestion, optimize routing transit routes, increase the efficiency of public transportation, and improve the overall transportation experience. By analysing real-time and historical data, these models offer solutions such as traffic rerouting, recommending efficient logistics routes, forecasting public transport demand, improve security, revolutionizing the way transportation is managed and transform urban mobility for the better.

MACHINE LEARNING MODELLING

The machine learning algorithm includes the following steps for analysis.

1) Data Collection: IoT devices collect data from various sensors, such as traffic cameras, GPS devices, motion, or light sensors. The data can be collected continuously or at regular intervals and transmitted to a central server or cloud platform for processing.

2) Preprocessing: Raw data collected from IoT devices often requires preprocessing to remove noise, handle missing values, or normalize the data. This step ensures that the data is in a suitable format for further analysis.

3) Feature Extraction:Feature extraction involves identifying and extracting relevant features from the pre-processed data. Features are specific measurements or characteristics that capture the essential information for the intended analysis or application. For example, in a smart commuting, features could include speed, Num_Reads, Hour, Zip Codes, Region, Bus Count.

4) Selection and Dimensionality Reduction: Depending on the application, it may be necessary to select a subset of features or reduce the dimensionality of the data. This step aims to eliminate irrelevant or redundant features, improving computational efficiency and reducing the risk of overfitting in machine learning models.

5) Data Integration: In some cases, data from multiple IoT devices or sources may need to be integrated to derive meaningful insights. Integration can involve combining data from various sensors, time synchronization, or merging data from different locations or devices.

6) Data Analytics: Once the relevant features have been extracted and processed, various analytics techniques can be applied to gain insights or make predictions. This can include statistical analysis, data mining, machine learning algorithms, or artificial intelligence models.

7) Visualization and Reporting: The processed data and analytics results can be visualized using charts, graphs, or dashboards to provide a clear representation of the information. Visualizations aid in understanding patterns, trends, or anomalies in the data. Additionally, reports or alerts can be generated to notify users or stakeholders of important findings or events.

8) Real-Time Processing: IoT systems often require real-time processing to enable timely decision-making or immediate actions based on the collected data. Real-time processing involves analysing data as it arrives and generating responses or triggers in near real-time.

9) Feedback Loop: The insights or actions derived from the processed data can be used to provide feedback and optimize the IoT system's performance. For example, adjusting sensor thresholds, improving predictive models, or triggering automated responses based on the analysis results.

Overall, feature extraction and data processing using IoT play a crucial role in transforming raw data collected from IoT devices into meaningful information and actionable insights for various applications such as smart homes, industrial monitoring, healthcare, or environmental monitoring and traffic management. Using Rank and percentile approach for the feature selection the following six features were found to be most relevant attributes for the data analysis and modelling.

Table 1 Rank and percentile

Sl. No.	**Attribute Selection**		
	Attribute Name	**Rank**	**Percent**
1	Speed	1	95.00
2	Num Reads	1	95.00
3	Hour	3	85.00
4	Zip Codes	3	85.00
5	Region	5	55.00
6	Bus Count	5	55.00

To analyse different prediction models, performance metrics such as TP Rate, FP Rate, precision and recall were used. These parameters are derived from a confusion matrix that shows the different ways in which the classification model gets confused when making predictions.

True Positive (TP) as in [11]refers to the number of predictions where the classifier correctly predicts the positive class as positive.

True Negative (TN)refers to the number of predictions where the classifier correctly predicts the negative class as negative.

False Positive (FP)refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.

False Negative (FN)refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

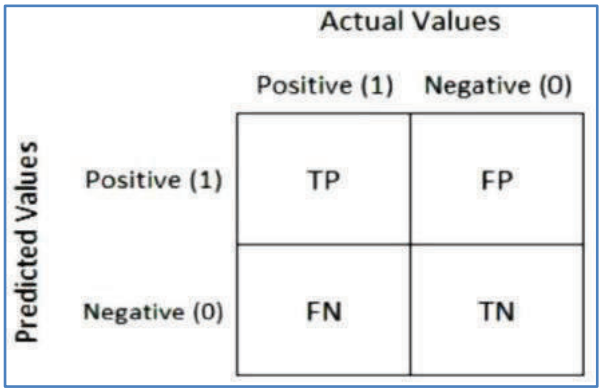


Figure 3 Confusion Matrix

ML PREDICTIVE MODEL ANALYSIS (LOW TRAFFIC)

Predictive machine learning models for smart transportation provide insights and data-driven solutions to solve complex urban mobility challenges. These models leverage large data sets from traffic sensors, GPS devices, and a variety of sources to predict traffic patterns, optimize routes, and improve public transportation systems. By leveraging real-time and historical data, they enable more efficient traffic management, reduce congestion, improve user experience, and promote sustainable transportation practices. Furthermore, these predictive models have the potential to play a central role in shaping future transportation policies and infrastructure developments for smarter and more accessible cities. The Weka tool is used to analyze Nine different classification machine learning algorithms. For the analysis of various predictive models, the performance measure like TP Rate, FP Rate, Precision and Recall were being used. Udaipur data set extracted from tomtom Server as in [12]was used with following features to analyze Machine Learning algorithms.

Table 2 Classifiers Performance Measure

Sl. No.	Class Label: Low Traffic				
	Classifier	**TP Rate**	**FP Rate**	**Precision**	**Recall**
1	Bayes Net	0.976	0.000	1.000	0.976
2	Naïve Bayes	0.971	0.060	0.996	0.971
3	Logistic	0.972	0.134	0.990	0.972
4	SM0	0.987	0.164	0.988	0.987
5	IBK	0.991	0.328	0.977	0.991
6	KStar	0.991	0.328	0.977	0.991
7	MultiClass Classifier	0.972	0.134	0.990	0.972
8	Random Forest	1.000	0.179	0.987	1.000
9	Random Tress	0.994	0.284	0.980	0.994

TP Rate as in[13]is used to measure the percentage of actual positives which are correctly identified by modelAccording to the performance measure TP rate for class label: Low Traffic it was found that the highest true positive rate was of the classifier Random Forest with value 1, followed by 0.994, 0.991 and 0.991 of Random Tree, IBK and KStar respectively whereas the lowest TP rate was found to be of the classifiers Naïve Bayes, Logistic and

Multiclass with values 0.971,0.972 and 0.972 respectively. Overall, it can be interpreted the most appropriate classifier based on the performance measure TP rate is found to be Random Forest. The "Random Forest" classifier appears to perform exceptionally well, with a TP Rate of 1.000, indicating perfect performance in distinguishing "Low Traffic" instances.

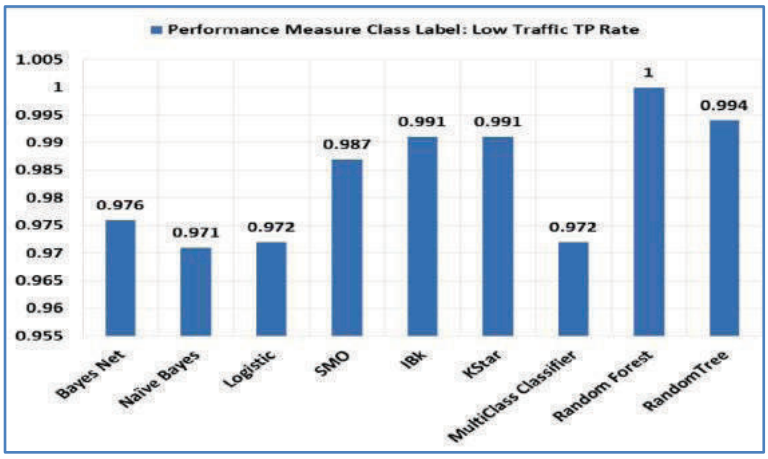


Figure 4 TP Rate for Class Label: Low Traffic

FP Rate also known as Type - I error is used to measure the percentage of actual positives which are incorrectly identified by model.Based on the performance measure FP rate for class label: Low Traffic it was found that the lowest false positive rate was of the classifier Bayes Net with value 0.000, followed by 0.06 of Navie Bayes whereas the highest FP rate was found to be of the classifiers IBK and KStar with values 0.328 each. Overall, it can be interpreted the most appropriate classifier based on the performance measure FP rate is found to be Bayes Net with lowest FP rate value.

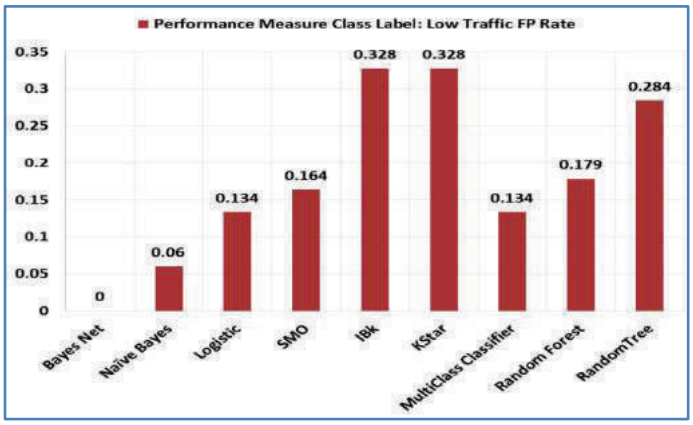


Figure 5 FP Rate for Class Label: Low Traffic

Precision as in [14] is the quality of a positive prediction made by the model. Precision refers to the number of True Positives divided by the total number of Positive predictions.

$$\text{Precision} = \frac{T_P}{T_P + F_P} \qquad (1)$$

According to the performance measure precision for class label: Low Traffic it was found that the highest precision value was of the classifier Bayes Net with value 1, followed by 0.996, 0.99 and 0.99 of Naïve Bayes, SMO and

Multiclass classifier respectively whereas the lowest precision value was found to be of the classifiers IBK and KStar with values 0.977 each. Overall, it can be interpreted the most appropriate classifier based on the performance measure precision is found to be Bayes Net.

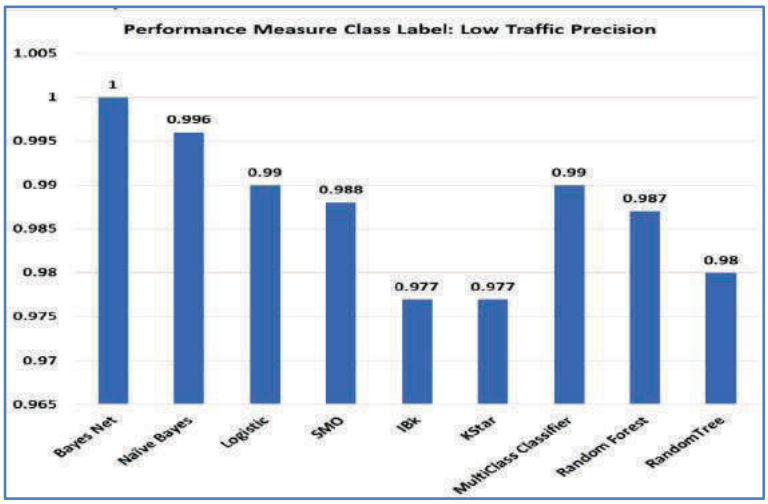


Figure 6 Precision for Class Label: Low Traffic

Recall as in[15]is the measures of how well a machine learning model can detect positive instances. It is also called as Sensitivity.Sensitivity refers to the number of true positives divided by the sum of True Positives and False Negatives. The model with high Sensitivity will have significantly fewer False Negatives.

$$\text{Sensitivity or Recall} = \frac{T_P}{T_P + F_N} \qquad (2)$$

Based on the performance measure recall for class label: Low Traffic it was found that the highest recall value was of the classifier Randon Forest with value 1.00, followed by 0.991 and 0.991 of IBK and KStar respectively whereas the lowest recall value was found to be of the classifiers Navie Bayes, Logistic and Multiclass with values 0.971, 0.972 and 0.972 respectively. Overall, it can be interpreted the most appropriate classifier based on the performance measure recall is found to be Random Forest.

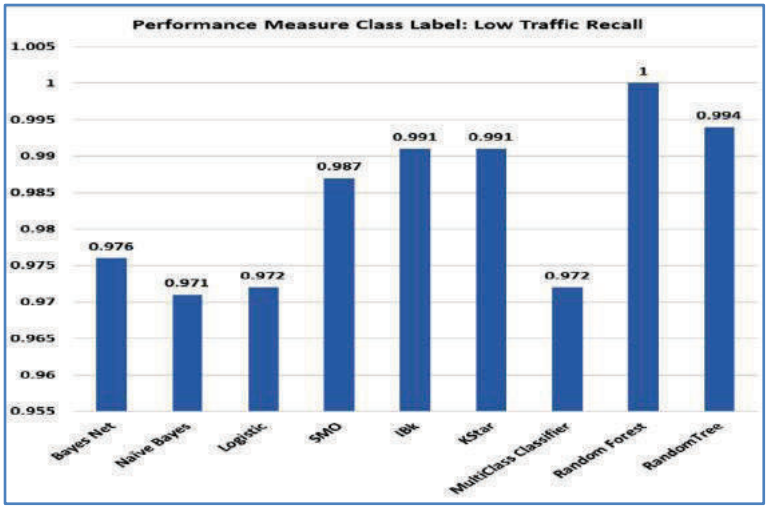


Figure 7 Recall for Class Label: Low Traffic

ML Predictive Model Analysis (High Traffic)

Sl. No.	Class Label: High Traffic				
	Classifier	**TP Rate**	**FP Rate**	**Precision**	**Recall**
1	Bayes Net	1.000	0.024	0.753	1.000
2	Naïve Bayes	0.940	0.029	0.700	0.940
3	Logistic	0.866	0.028	0.690	0.866
4	SM0	0.836	0.013	0.824	0.836
5	IBK	0.672	0.009	0.849	0.672
6	KStar	0.672	0.009	0.849	0.672
7	MultiClass Classifier	0.866	0.028	0.690	0.866
8	Random Forest	0.821	0.000	1.000	0.821
9	Random Tress	0.716	0.006	0.716	

According to the performance measure TP rate for class label: Heavy Traffic it was found that the highest true positive rate was of the classifier Bayes Net with value 1.0, followed by 0.94 , 0.866 and 0.866 of Naïve Bayes, Logistic and Multiclass respectively whereas the lowest TP rate was found to be of the classifiers IBK and KStar with values 0.672 each. Overall, it can be interpreted the most appropriate classifier based on the performance measure TP rate is Bayes Net.

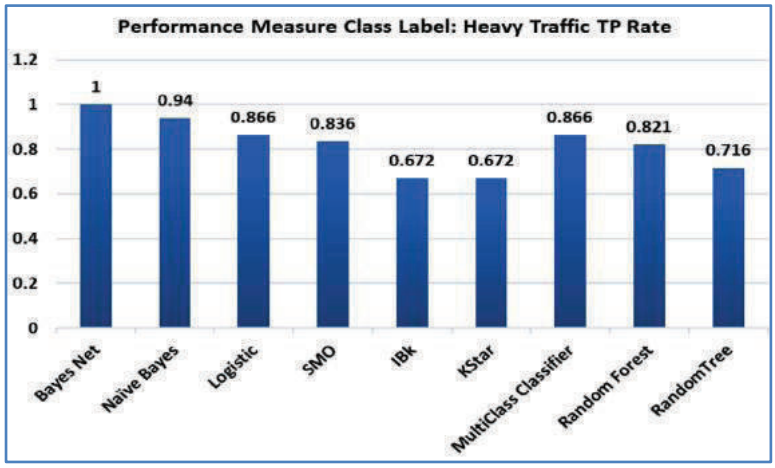


Figure 8 TP Rate for Class Label: Heavy Traffic

Based on the performance measure FP rate for class label: Heavy Traffic it was found that the lowest false positive rate was of the classifier Random Forest with value 0.00, followed by 0.006 of Random Tree whereas the highest FP rate was found to be of the classifiers Naïve Bayes, Logistic and Multiclass with values 0.029, 0.028 and 0.028 respectively. Overall, it can be interpreted the most appropriate classifier based on the performance measure FP rate is found to be Random Forest with lowest FP rate value.

According to the performance measure precision class label: Heavy Traffic it was found that the highest precision value was of the classifier Randon Forest with value 1.0, followed by 0.889, 0.849 and 0.849 of Random Tree, IBK and KStar respectively whereas the lowest precision value was found to be of the classifiers Logistic, Multiclass and Naïve Bayes with values 0.69, 0.69 and 0.7 respectively. Overall, it can be interpreted the most appropriate classifier based on the performance measure precision is found to be Random Forest.

Based on the performance measure recall class label: Heavy Traffic it was found that the highest recall value was of the classifier Bayes Net with value 1.0, followed by 0.94 , 0.866 and 0.866 of Naïve Bayes, Logistic and Multiclass respectively whereas the lowest recall value was found to be of the classifiers IBK and KStar with values 0.821

each. Overall, it can be interpreted the most appropriate classifier based on the performance measure recall is found to be Bayes Net.

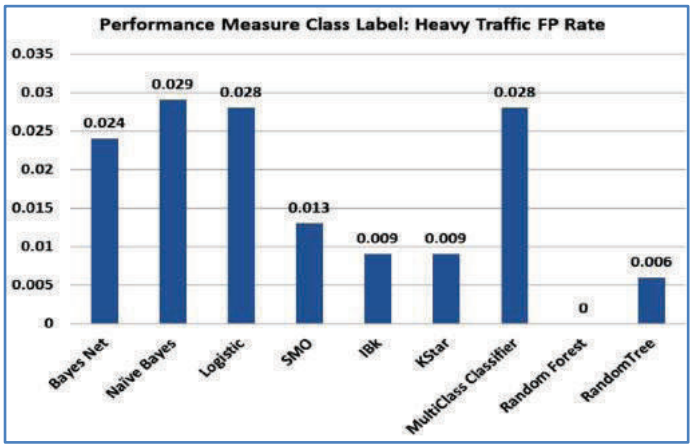


Figure 9 FP Rate for Class Label: Heavy Traffic

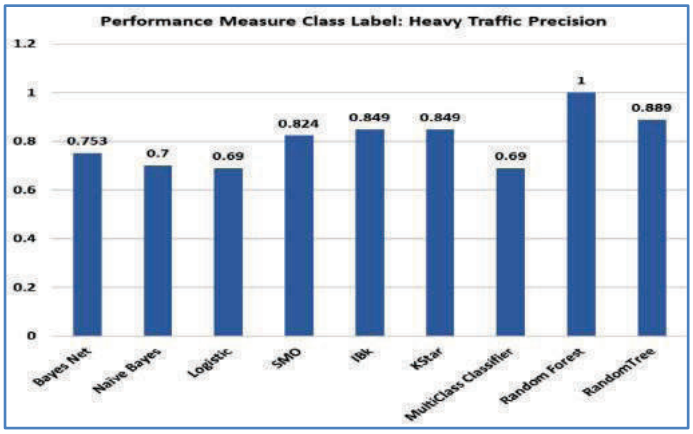


Figure 10 Precision for Class Label: Heavy Traffic

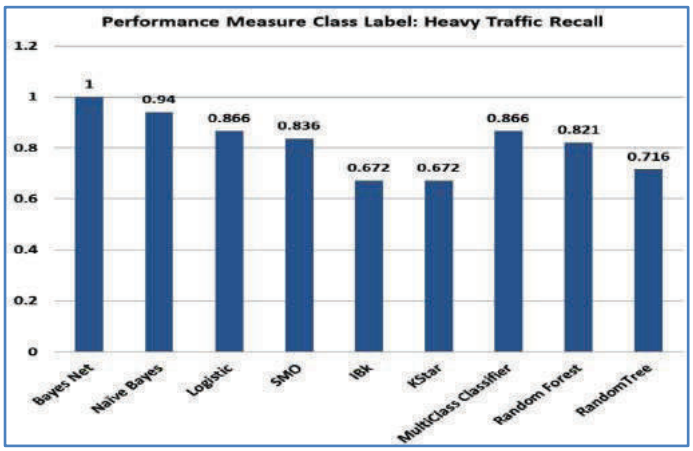


Figure 11 Recall for Class Label: Heavy Traffic

SUMMARY

The major findings related to the comparative analysis of machine learning predictive models using Weighted Sum Model (WSM) approach is shown below.

1. Random Forest is the best and most appropriate classifier for traffic congestion control and traffic flow as it is having the highest score of 0.98.

2. Bayes Net is the second most appropriate algorithm with score of 0.6530

3. The third best classifier being identified is SMO with score of 0.4871.

For predicting the Udaipur traffic flow Random Forest is the most appropriate algorithm.

Table 3 Weighted sum model

Classification Models	Score (Weighted Sum)
Random Forest	0.9800
Bayes Net	0.6530
SMO	0.4871
Random Tree	0.4110
Naïve Bayes	0.3111
KStar	0.1745
Logistic	0.1406
MultiClass Classifier	0.1406
IBK	0.0679

CONCLUSION AND FUTURE WORK

Machine learning algorithms provide powerful capabilities for a variety of tasks when provided with high-quality, validated data. Maintaining model accuracy over time requires continuous model monitoring and adjustment. This approach includes real-time data analysis, model retraining as data evolves, and automatic deployment of updated models to ensure sustained performance and relevancy. A K-fold cross-validation approach is required. A fold is a set of records in a dataset, and k is the number of folds that affects the performance and reliability of the model. In general, higher values of k increase variance and computational cost, but decrease bias.Furthermore, Data Analysis with class label of different traffic densities apart from Low Traffic and high traffic will increase the accuracy and reliability of the model.

REFERENCES

1. M. King, B. Zhu, and S. Tang, “Optimal path planning,” Mobile Robots, vol. 8, no. 2, pp. 520-531, March2001.
2. Deak, G., & Walravens, N. (2019). The future of urban mobility: Towards integrated smart mobility solutions. Transport Reviews, 39(2), 157-181.
3. Erlmann, S., & Dantzig, S. V. (2018). Exploring the potential of shared mobility services to alleviate urban traffic congestion: A literature review. Transport Reviews, 38(6), 769-793.
4. Fintikakis, N., & Bourka, A. (2018). Commuting patterns and smart cities: A comparative study of Stockholm and London. TeMA Journal of Land Use, Mobility and Environment, 11(3), 239-254.
5. Gkiotsalitis, K., & Cats, O. (2018). Assessing the impact of autonomous vehicles on our cities. Transportation Research Part A: Policy and Practice, 118, 407-417.
6. Gohar, M., Sagheer, A., Javaid, N., Anpalagan, A., & Khan, F. (2019). Fog computing-enabled intelligent transportation system: Architecture, opportunities, and challenges. IEEE Access, 7, 29037-29057.
7. Hall, R. E., & Pfeiffer, U. (2016). The impact of smart cities on travel behavior: An exploratory study. Transportation Research Part A: Policy and Practice, 88, 96-114.

8. Jung, C., & Couclelis, H. (2017). Smart cities and the politics of urban data. Journal of Planning Education and Research, 37(4), 457-466.
9. Kaur, A., & Kaur, P. (2017). Internet of Things for transportation: A review. IEEE Sensors Journal, 17(14), 4371-4378.
10. Khan, A., Salah, K., & Zeadally, S. (2019). Fog computing for intelligent transportation systems: Opportunities and challenges. Computer Networks, 160, 96-112.
11. https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826
12. https://www.tomtom.com/products/traffic-stats/
13. https://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7_255#:~:text=Definition,positives%20which%20are%20correctly%20identified.
14. https://c3.ai/glossary/machine-learning/precision/#:~:text=Precision%20is%20one%20indicator%20of,the%20number%20of%20false%20positives).
15. https://www.shiksha.com/online-courses/articles/sensitivity-vs-specificity/#:~:text=Sensitivity%20measures%20how%20well%20a,have%20significantly%20fewer%20False%20Negatives.

Springer

ICMCSI

Presentation Certificate

This certificate is awarded to

Dangwani Avinash

have successfully presented the paper entitled

Use of AI in Cloud based Certificate Authentication for Travel Concession.

at the

4th International Conference on
"Mobile Computing and Sustainable Informatics"
(ICMCSI 2023)

organized by
Tribhuvan University, Nepal on 11th & 12th January, 2023

Session Chair

Organizing Secretary
Dr. S. Smys

Conference Chair
Prof. Dr. Subarna Shakya

Proceedings

of the

4th International Conference on Mobile Computing and Sustainable Informatics

ICMCSI 2023

11-12, January 2023 | Lalitpur, Nepal

Springer

organized by

Tribhuvan University

Nepal

14	The Application of Mobile Phones to Enable Traffic Flow Optimisation *T Shilowa, J P van Deventer, M J Hattingh*
15	Authentication in Cloud Computing: Open Problems *Zarif Khudoykulov, Abdukodir Karimov, Ravshan Abdurakhmanov, Mirkomil Mirzabekov*
16	A Hypothesis on Cloud Sourcing Sharing Users' Mobile Devices through Virtualization *Nazmus Sakib, Al Hasib Mahamud*
17	AN ANALYTICAL ASSESSMENT OF MACHINE LEARNING ALGORITHMS FOR PREDICTING CAMPUS PLACEMENTS *S Bala Dhanalakshmi, Raksheka Rajakumar, Swetha Shankar, R Sowndharya Rani, N Deepa, C Subha Priyadharshini*
18	A Systematic Literature Review on Symmetric and Asymmetric Encryption Comparison Key Size *Mohammed Althamir, Abdullah Alabdulhay, Muhammad M Yasin*
19	Building Hindi Text Dataset on Stock Market Tweets and Sentiment Analysis Using NLP *Anushka Choudhary, Mohit Gupta, Lavanya S K*
20	Fine Tuning of RoBERTa for Document Classification of Arxiv Dataset *Kshetraphal Bohara, Aman Shakya, Bishal Debb Pande*
21	Comparative Analysis of Using Event Sourcing Approach in Web Application Based on the LAMP stack *Marian Slabinoha, Stepan Melnychuk, Vitalia Kropyvnytska, Bohdan Pashkovskyi*
22	Profitability Improvement for CV. XYZ with Improvement of Work Posture and Workflow Using REBA and Modelling Simulation *Louis Valentino, Lina Gozali, Frans Jusuf Daywin, Ariawan Gunadi*
23	Demand Forecasting Using Time Series and ANN with Inventory Control to Reduce Bullwhip Effect on Home Appliances Electronics Distributors *Stiven Tjen, Lina Gozali, Helena Juliana Kristina, Ariawan Gunadi, Agustinus Purna Irawan*
24	Memory Malware Identification via Machine Learning *Maysa Khalil, Qasem Abu Al-Haija*
25	Basketball Shot Conversion Prediction using various ML techniques and its analysis *Sanyam Raina, Shreedhar Bhatt, Vaidehi Shah, Heem Amin, Vinay Khilwani, Dr. Samir Patel*
26	Progressive Web App Implementation in Omah Wayang Klaten Website *Budi Susanto, Gloria Virginia, Umi Proboyekti, Jeysy Carmila Dewi Ester*
27	Use of AI in cloud based certificate authentication for travel concession. *Dangwani Avinash, Jetawat Ashok Kumar, Rawat Chandansingh*
28	Media Player Controller using Hand Gestures *Vijay Mane, Harshal Baru, Abhishek Kashid, Prasanna Kshirsagar, Aniket Kulkarni, Prathamesh Londe*

Use of AI in cloud based certificate authentication for travel concession.

* Dangwani Avinash[1], Jetawat Ashok kumar[2], and Rawat Chandansingh[3]

[1] Computer Department Pacific Hills, Pratap Nagar Extension, Airport Road, Debari, Udaipur—313024, Rajasthan, India
[1*]avin861@gmail.com

[2]Computer Department Pacific Hills, Pratap Nagar Extension, Airport Road, Debari, Udaipur—313024,Rajasthan, India
[2] drashokjain61@gmail.com

[3]Vivekanand Education society institute of technology collectors colony, Mumbai– 400074, Maharashtra, India
[3] chandansingh.rawat@ves.ac.in

Abstract. Certificate Authentication is a big challenging task in socio economic country like India. 28 states and 8 union territories with diversified cultures and languages makes it big voluminous task. There is urgent need to automate this authentication task. This paper proposes the cloud based certificate authentication. Google cloud services are used to automate authentication process. Vision API and flask framework is explored which allows developers to easily integrate vision detection features within applications including Image labeling, Face and landmark detection, Optical character recognition , Tagging of explicit contents. The proposed arrangement make use Cloud Vision API optical character recognition to infer presence of required fields in scanned pdf. Recognized fields will be communicated in the output report.

Keywords: Caste Authentication, Google cloud vision, Open computer vision, Optical character recognition, APISetu, Google Cloud Vision, Application program interface, Cloud bucket.

1 Literature review

H. Gaikwad, N. D'Souza, R. Gupta and A. K. Tripathy(2021) found millions of students every year go through lengthy and cumbersome process of document verification for their higher studies. This results in significant overhead as documents are transferred between institutions for verification. There is a need for an automated credential verification system which can reduce the time required for the document verification process. They have used Blockchain Technology that can be used to reduce overhead and reduce the time taken for document verification

from days to mere seconds.

Jignasha Dalal Meenaland Chaturvedi Himani Gandre and Sanjana Thombare(2020) proposed solution of biometric to access all the previous degree certificates of a students using blockchain. The students will submit the hash of their biometric and a unique phrase. This hash will be stored on the blockchain. The degree certificate for a student will be issued by the college authorities. They will upload the hash of digitally signed certificate on the blockchain. They proposed to link documents with person's identity without involving a third party.

D. Vaithiyanathan and M. Muniraj(2019) has worked on "Cloud based Text extraction using Google Cloud Vison for Visually Impaired applications". They have designed assistive device called smart reader that is capable of capturing an image from a camera and extract the text from the captured image. Text is converted to speech as voice based output to assist the visually impaired people.

2 Introduction

This paper will emphasize on Google services and it's API for developing cloud based certificate authentication system. Manual authentication of certificates for senior citizens, students & citizens belonging to certain gender and community becomes sometimes confusing and time consuming. The inspector who validates documents for giving travel concession needs substantial amount of proof and supporting documents to authenticate certificates. Citizens specially senior citizens sometimes have to face stressful situation due to delayed and cumbersome authentication process. Recently Maharashtra state in India has given complete state transport travel free to senior citizens above 75 years. Our proposed solution consists of a cloud web app which will have a front-end to get applicant details, certificate submission and requesting authentication, along with a Google cloud backend which will have three modules: An OCR API module to extract applicant details from certificates, An compare module to compare extracted details with submitted details and a report module to send and display authenticated data stored in the cloud storage. Google cloud is acting as third party trusted centralised authority which provides software as service and centralized storage for authentication of certificates against distributed systems which uses blockchain technology to authenticate academic certificates[1].

With the advancement in the field of computer vision Artificial intelligence it is possible to extract text using OCR[2]. Pattern recognition technique in optical character recognition provides accurate results in extracting text from the various document formats such as JPG & PDF. Automation of authentication process reduces the administrative overhead by minimizing the use of paper and curtailing

the verification process. It provides a Real-Time verification module enabling agencies to verify data directly from issuers after obtaining user consent[3-10].

The proposed certificate authentication model for travel concession is Organized into following sections, Authentication System, System Design, Experimental Results, Summary, Conclusion and Future Work.

3 Authentication System

The main aim is to develop quick and easy to use cloud based certificate authentication system which uses centralize cloud SQL database to compare OCR API extracted features with the stored features collected from user with the help of front end designed using HTML and JavaScript. Google cloud vision OCR Functions are used to recognize the text in pdf document and are used to convert pdf image into accessible electronic text[10-17].

To overcome the drawback of fraudulent intentions of applicant and forged documents two level Authentication system Local Level and Government Server level is proposed as in Fig 1.

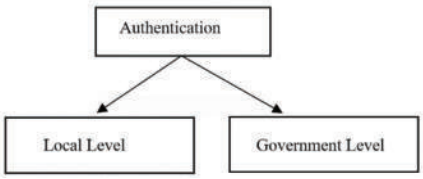


Fig. 1. Authentication types

3.1 Local Level Authentication

For Local level authentication OCR extracted text is compared with the stored cloud SQL database. Cloud SQL is Fully managed relational database service for MySQL, PostgreSQL, and SQL Server with rich extension collections, configuration flags, and developer ecosystems. Cloud SQL has many advantages like reduced maintenance cost, fully reliable and secured 24/7 service, server instances can be scaled effortlessly when demand increases.

3.2 Government Level Authentication

After successful comparison of all required attributes second level is proposed which has some government restrictions like GST Number. For Government server level APISetu or uidia.gov.in site can be used. APISetu provides single platform access to information from multiple sources. It can be used for a variety of use cases such Know Your Client (KYC) and other authentication services.

4 System Design

The cloud web application program is designed to authenticate local fields like Certificate type, name of applicant, state, Aadhaar card, Date of Birth, Age & School name. Front end is designed in HTML and backend code is written in python and flask framework.

Google Cloud vision API technology identify the content of an image with the help of powerful machine learning models. REST API in Google Cloud Vision API enables developers to understand the content of an image by encapsulating powerful machine learning models. It has powerful pre-trained machine learning models. It offers powerful image analysis, insight from your images, detect and classify multiple objects including the location of each object within the image. & detect required content. It can read printed and handwritten text, and build valuable metadata into your image Cloud vision API can be used with Auto ML Vision to automate the training of your own custom machine learning models. These models can be optimized for accuracy, latency and size.

Google Cloud vision AI is designed to understand text with pre trained vision API models. The research motivation and the will to explore more about Cloud vision API is created after knowing the applications of the big companies like New York Times and box. New York Times is using Google cloud to preserve the visual history by finding out many untold stories in millions of archived photo's & Box company is using image recognition and OCR API of Google Cloud vision for content management. Cloud vision technology has enabled New York Times to unveil more than a century of global events that have shaped our modern world. The New York Times built a processing pipeline that stores and processes the photos and will use cloud technology to process and recognize text, handwriting and other details that can be found in the images. Box company extracted printed words from the scanned image and then returned labels and recognized characters in JSON responses. Fig. 2 shows the System Design.

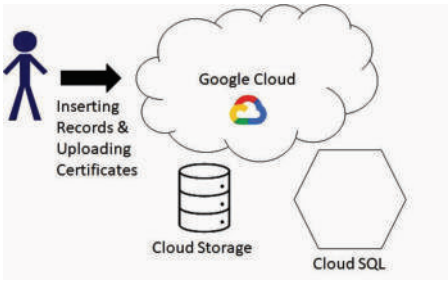


Fig. 2. Uploading credentials

Since certificate authentication is most widely used for students bus fare concession, Senior citizen bus fare in india therefore sample student certificates, Aadhaar card

and senior citizen Aadhaar card were selected for developed cloud based software testing. Another area where certificate authentication is used for students belonging to Schedule caste, Schedule Tribe and Other backward class category in India.

Front end is designed using HTML forms to get required details / records from the student. All these records are maintained in cloud SQL tables. Cloud SQL is fully managed relational database service for MySQL, PostgreSQL, and SQL Server with rich extension collections, configuration flags and developer ecosystems. It provides reliable, secure and scalable solution for cloud data base which ensure that operations should run 24 * 7 for 365 days without any disruptions. Cloud SQL automates all your backups, replication, encryption patches, and capacity increases—while ensuring greater than 99.95% availability, anywhere in the world.

Three cloud SQL Tables, Testify table, Reference Table and Flag Table were created using SQL CREATE TABLE Query command and used for HTML form backend storage.

4.1 Testify Table

Table 1. Testify table

Name	Email id	State	Aadhar Card Number	Date of Birth	Age	School Name

Testify_table stores personal details of applicant like name, Email id, state, Unique identification number (Aadhaar card), Date of Birth, Age, School name.

Unique id Aadhaar Card Number is the key which is used to search particular record in Testify_table. To make system more easy to use and user friendly, another search options are also included. Different ways SQL data can be searched from Testify_table are, Search by Aadhaar Card Number , Search by Name , Search by Email id and Search by School Name.

4.2 Reference Table

Table 2. Reference table

Aadhar Card Number	File_Name	Certificate type

Reference table was created with the aim to provide reference for caste certificate file storage in Google Cloud storage buckets. Reference table helps designed code

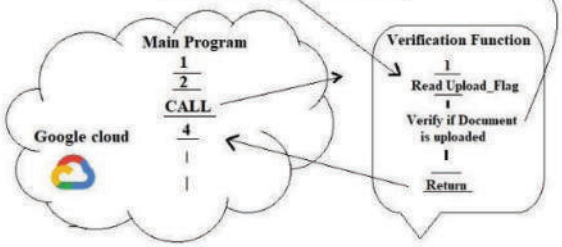
to understand which certificates are uploaded in backend. The Null entry in Certificate type column indicates that file is not uploaded in cloud bucket. Code can be designed to send email message “Certificate type not uploaded” to applicant.

4.3 Flag Table

Table. 3 Flag Table

Flag_Table

Document Upload_Flag	Document Verified_Flag
0	0
1	0
1	0
1	1 (Verified)



Flag table was created to mark different conditions of file upload to know whether required caste certificates for authentication are loaded or not. Flags in flag table are basically sequence of pre-defined binary bits which holds true / logic 1 when required certificates are uploaded and false / logic 0 when required certificates are not available in cloud storage bucket.

4.4 Process Model

Process flow is depicted in Fig.3 When user types URL on address bar of Browser, it get’s designed HTML index Form for inputting data like Name, Email id, state, Aadhaar card Number, Date of Birth, Age, School name. All these details are loaded in Testify_table (Table 1). Upon pressing next button in HTML index Form it will display another form to upload required scanned pdf files like Aadhaar card & Bus ticket conession application form. All these pdf files will be loaded in Google cloud bucket which is created in Google cloud storage. Once the files will be loaded in Google cloud bucket then reference table (Table 2) will be updated with Aadhaar card Number, File name prefixed by Aadhaar card Number & Certificate type. To know whether applicant has uploaded required document or not flag_table is updated with logic 1 for particular document. Logic is shown in (Table 3). Main Program in Google cloud will read Upload_Flag for particular certificate from Flag_Table and will invoke OCR API code for reading Certificates.

OCR API has async_batch_annotate_files() annotation function which detect text and image for a batch of generic files, such as PDF documents at once. This function can be invoked in python language by importing vision module from google cloud which will generate unstructured data which is stored in Blob list. Blob is object storage solution for the cloud. It is optimized for storing massive amounts of unstructured data such as text or image binary data. Blob data is converted into JSON string data, which is finally converted into text format to search required text field is present or not.

Training code is required to train designed software about the format used in the uploaded document. Execute Verify function to find required text field is available as per trained model. If required text (name, state, city, date of birth, age etc) is found then document verified_flag will be updated to logic 1 and output report is generated where the verified column holds identified / not identied condition.The fields identified will be notified in “Fields verified” column.

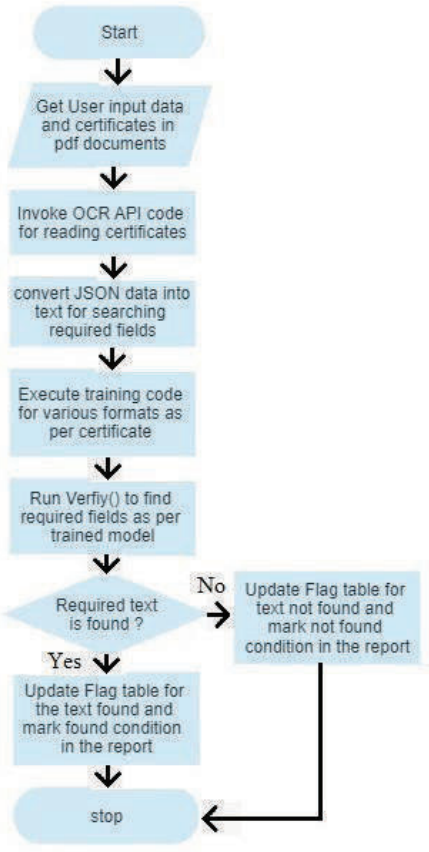


Fig. 3. Design Flow chart

5 Experimental Results

5.1 Front end Dashboard

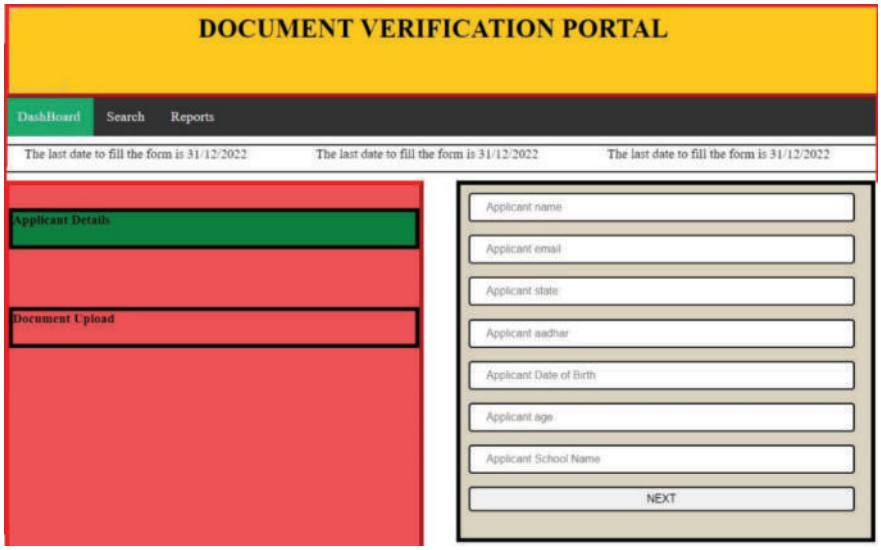


Fig. 4. Front end

Snap shot of front end is given in fig 4. The objective of front end is pleasent and convenient user interface. Front end has three menu options, Dashboard, Search and Reports. Dashboard has designed HTML form to take applicant details and tab for Document Upload. All entries are recieved from user are loaded in SQL testify_table of SQL instance data base as shown in Table 4.

Table. 4 Flag Table

Name	Email_id	State	Aadhaar_Card_Number	Date_of_Birth	Age	School_Name
Romil	Romil_raj@gmail.com	Goa	790195155740	26/10/2004	18	BITS Pilani
Riya	Riya2005@gmail.com	Kerala	527295155654	01/11/2005	17	St.Joseph

After Applicant details are entered, user is prompted to upload Aadhaar card and concession form details as shown in fig. 5. Due to functional limitations we have restricted file upload to pdf only. With modification other files format upload can also be allowed. After uploading, documents are stored in Google cloud bucket and designed code will update Reference table with the details as shown in Table 5.

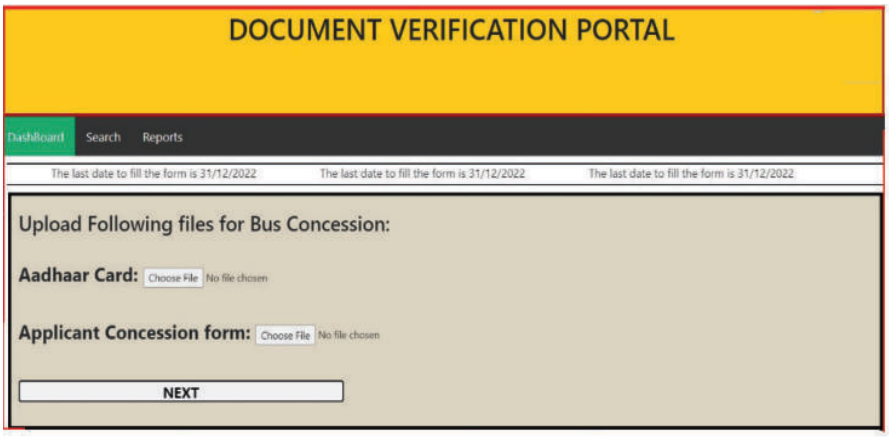


Fig. 5. Document upload

Table. 5 Flag Table

Aadhaar_card_Number	File_name	Certificate_type
790195155740	790195155740_Aadhaarcard.pdf	Aadhaarcard
527295155654	527295155654_Concessionform	Concessionform

5.2 Front end Search

Effective search always improve productivity, enhances decision making and makes management easier. Fig. 6. Shows search options.

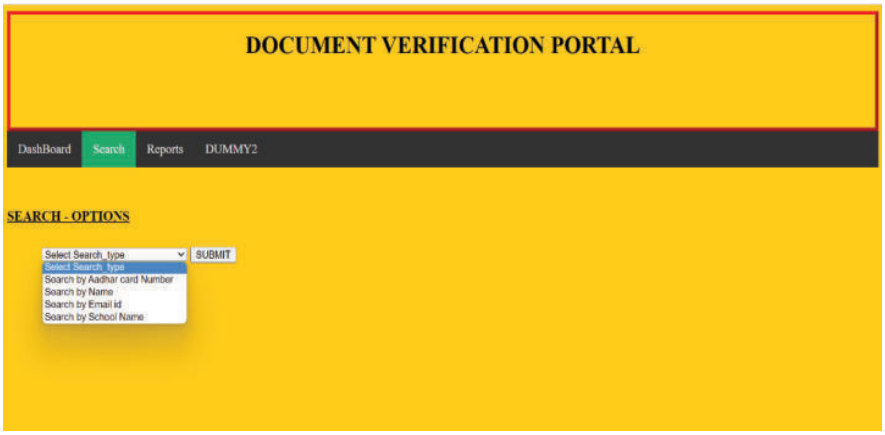


Fig. 6. Search options

Different search options enable administrator to search records by Aadhaar card Number (Unique identification number), Search by Name, Search by Email id and Search by School Name.

5.3 Front end Reports

Reports gives you real time information at your finger tips and improves speed in decision making. Efficient reports always save allot of time and makes management easier. If required fields in certificate are not authenticated or wrong certificates are uploaded then software is generating output shown in fig. 7. Output for correct certificate upload is shown in fig. 8.

Certificate Type	Verified	Fields Verified
Aadhaar Card	✗	Name , Aadhaar Card Number State
Student Concession Application Form	✗	Name, State, Age , School name

Fig. 7. Report with wrong certificate upload

Certificate Type	Verified	Fields Verified
Aadhaar Card	✓	Name , Aadhaar Card Number State
Student Concession Application Form	✓	Name, State, Age , School name

Fig. 8. Report with valid certificate upload

6 Summary

To overcome misery of slow and manual authentication work and to keep pace with emerging trends and innovation this project was selected. This project has explored the world of artificial intelligence to uplift society by revolutionizing smart work culture in society.

There was big question in front of us whether to develop independent PC Application or to develop cloud based software. Looking at several merits like scalability, availability, advance security, data loss prevention and collaborative work environment the cloud project was selected. There are four major cloud service providers in the world, Microsoft AZURE, Amazon Web Services(AWS), IBM cloud services and google cloud platform (GCP). Google cloud platform was selected due to its economical charges and localised services.

Artificial intelligence computer vision technology has opened doors for various innovations, image to text conversion using OCR API is one of them. various functions are developed in cloud technology to convert single or group of images into text. Various formats like PDF and JPEG are converted into JSON formats and finally into text formats for editing, this text can be used to compare with required fields for getting desired results. Lakhs of images can be converted and compared in very less time using cloud technology, which is practically not possible for human being.

7 Conclusion and Future Work

We have investigated use of Google cloud vision AI - API for certificate authentication. This cloud automation will simplify authentication process of concession issuing authority up to great extent and speed up process of issuing concession in smart cities. The main practical limitation of this project is second level authentication, which is not possible without the help of Government servers. With the help of Govt API available on site https://www.uidai.gov.in/914-developer-section.html we can opt for second level authentication also in future. Face detect feature of Google cloud vision AI – API can be explored for better results in future. Allot of work is in progress in cloud vision API and cloud technologies. Government of India vision for NRC (National register for citizens) will speed up cloud authentication and automation.

References

[1] H. Gaikwad, N. D'Souza, R. Gupta and A. K. Tripathy, "A Blockchain-Based Verification System for Academic Certificates," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), 2021, pp. 1-6, doi: 10.1109/ICSCAN53069.2021.9526377.

[2] D. Vaithiyanathan and M. Muniraj, "Cloud based Text extraction using Google Cloud Vison for Visually Impaired applications," 2019 11th International Conference on Advanced Computing (ICoAC), 2019, pp. 90-96, doi: 10.1109/ICoAC48765.2019.246822.

[3] Q. Liu Q. Guan X. Yang H. Zhu G. Green and S. Yin "Education-Industry Cooperative System Based on Blockchain" 1st IEEE International Conference on Hot Information-Centric Networking (HotICN) pp. 207-211 15–17 August 2018.

[4] Jignasha Dalal Meenaland Chaturvedi Himani Gandre and Sanjana Thombare "Verification of Identity and Educational Certificates of Students Using Biometric and Blockchain (April 8 2020)" Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST) 2020.

[5] Yasuhisa Fujii "Optical Character Recognition Research at Google" IEEE 7th Global Conference on Consumer Electronics (GCCE) December 2018.

[6] Heba Saleous Anza Shaikh Ragini Gupta and Assim Sagahyroon "Read2Me: A cloud-based reading aid for the visually impaired" International Conference on Industrial Informatics and Computer Systems (CIICS) May 2016.

[7] Digilocker facility provided by Government of india https://www.livemint.com/money/personal-finance/all-you-need-to-know-about-digilocker-and-how-to-use-it-11612943898102.html

[8] Blockchain vs cloud comparison https://www.upgrad.com/blog/blockchain-vs-cloud-computing/

[9] Cloud computing vs distribute computing https://www.projectpro.io/article/cloud-computing-vs-distributed-computing/94

[10] Use of Python for fetching data from database. https://towardsdatascience.com/pull-and-analyze-financial-data-using-a-simple-python-package-83e47759c4a7

[11] Atlas platform for paperless KYC. https://documenter.getpostman.com/view/12409759/TVCZaWzp#intro

[12] Postman use with google cloud platform https://pnatraj.medium.com/google-cloud-api-with-postman-f4cf070e665f

[13] Unique Identification Authority of India https://www.uidai.gov.in/ecosystem/authentication-devices-documents/developer-section/916-developer-section/data-and-downloads-section.html

[14] Government of India Online API Library. https://apisetu.gov.in/api/directory

[15] JSON Formatter site. https://chrome.google.com/webstore/detail/json-formatter/bcjindcccaagfpapjjmafapmmgkkhgoa?hl=en

[16] Google cloud vision API https://stackshare.io/stackups/google-cloud-vision-api-vs-opencv

[17] Maharashtra state transport concession form. https://msrtcblog.blogspot.com/2017/10/download-msrtc-student-concession-form.html.

12th INTERNATIONAL CONFERENCE

CERTIFICATE OF PARTICIPATION

16-17 April 2021

SUSTAINABLE GLOBAL TRENDS : PLANET, PEOPLE AND PROFIT

PACIFIC INSTITUTE OF MANAGEMENT

PACIFIC ACADEMY OF HIGHER EDUCATION AND RESEARCH UNIVERSITY, UDAIPUR

&

PACIFIC BUSINESS SCHOOL

RAJASTHAN TECHNICAL UNIVERSITY, KOTA

This is to certify that **Mr. Avinash Dangwani**

of PAHER University, Udaipur (Research Scholar)

has presented his/her paper entitled Data Analytics Sales Prediction Model

during the two-day International Conference jointly organized by Pacific Institute of Management & Pacific Business School, Udaipur in online mode. The Organizers wish his/her success in future endeavors.

Prof Mahima Birla	Prof Dipin Mathur	Dr Shivoham Singh
Conference Chair	Director Conference	Organizing Secretary

ISSN: 0974-438X
Impact factor (SJIF):8.396
RNI No.:RAJENG/2016/70346
Postal Reg. No.: RJ/UD/29-136/2017-2019

Pacific Business Review International

A Refereed Monthly International Journal of Management Indexed With Web of Science(ESCI)

Editorial Board

Editorial Team

A Refereed Monthly International Journal of Management

August 2021

Name	:	Index
Download	:	

Name	:	The Effect of Product Photograph and Information on Digital Apparel Marketing
Author	:	Ahmet Özbek, Cansu Tor-Kadıoğlu, Ahasanul Haque

Already have a manuscript?

Use our Manuscript Matcher to find the best relevant journals!

Find a Match

Refine Your Search Results

0974-438X

Search

Sort By: Title (A-Z)

Filters

Clear All

- *Web of Science* Coverage
- Open Access
- Category
- Country / Region
- Language
- Frequency

Search Results

Found 1 results (Page 1)

Share These Results

Exact Match Found

PACIFIC BUSINESS REVIEW INTERNATIONAL

Publisher: PACIFIC INST MANAGEMENT , PACIFIC HILLS, PRATAP NAGAR EXTENSION, AIR PORT RD, UDAIPUR, RAJASTHAN, INDIA, 313 001

ISSN / eISSN: 0974-438X

Web of Science Core Collection: Emerging Sources Citation Index

Share This Journal

View profile page

* Requires free login.

Data Analytics Sales Prediction Model

Avinash Dangwani
PhD Scholar
Department of Engineering (Computers)
Pacific University
Airport Road, Debari, Udaipur (Rajhasthan)

Dr. Chandansingh Rawat
Associate Professor
Dept of Electronics & Telecommunication
Engineering VESIT HAMC Collectors Colony,
Chembur Mumbai

Abstract

In every New financial year Company propose Advertisement Budget to improve their sales. Estimation of Advertisement Budget is not easy task as it involves financial parameters. Managers are always interested to know prediction model for sales which is function of Advertisement expenses.

This paper will develop Sales prediction model using simple linear regression. The model will be built using the training dataset to estimate the regression parameters. The method of Ordinary Least Squares (OLS) will be used to estimate the regression parameters using python. Regression model will be validated to ensure goodness of fit before it can be used for practical application. The single variable regression is the limitation of this model. In future multiple variables can be calculated using multiple linear regression model using python.

Keywords:

Simple linear regression, Ordinary Least Square (OLS), Training & validation data, Sum square regression (SST), Total sum of squares (SSR).

Introduction

This paper will develop sales prediction model using Simple Linear regression. sales prediction has two main methods(1) Qualitative method, (2) Quantitative method [3].Some of the Qualitative methods are Expert's Opinion Method, Sales Force Composite Method, Survey of Buyer's Expectations, Historical Analogy Method, Jury of Executive Opinions & Leading Indicators Method.

Some of the Quantitative methods areTest Marketing, Time Series Analysis, Moving Average Method, Exponential Smoothing Method, Regression Analysis&Econometric Models.

This paper will explore sales prediction using regression analysis due to its lower time complexity as compare to some of the other algorithm, Furthermore, these models can be trained easily and efficiently even on systems with relatively low computational power when compared to other complex algorithms. Building a regression model is an iterative process and several iterations may be required before finalizing the appropriate model [2]. Regression model is Organized in following sections.

ØSection – I: Simple Linear regression

ØSection – II: Ordinary least square(OLS) Method.

ØSection – III: Results& Model Diagnostics.

ØSection – IV: Conclusion

Simple Linear Regression

Simple linear regression (SLR) is a statistical technique which uses the existence of an association relationship between a dependent variable (outcome variable) and an independent variable(predictor/feature variable).

The functional form of SLR is as follows

$Yi = \beta0 + \beta1\ Xi + \varepsilon i$ (1)

Where

Yi =Value of the ith observation of the dependent variable

Xi = Independent variable of ith observation

εi = random error (residuals) in predicting thevalue of Yi

β0 & β1 = regression parameters

Ordinary least square (OLS) Method

Equation (1) can be re written as

$\varepsilon i = Yi - \beta0 - \beta1\ Xi$ (2)

The regression parameters β0& β1 are estimated by minimizing the sum of squared errors(SSE)

$$SSE = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i) \quad (3)$$

The estimated values of regression parameters are given by taking partial derivative of SSE with respect to β_0 & β_1 and solving resulting equation for the regression parameters. The estimated parameters are given by

$$\hat{\beta}_1 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y_i - \bar{y})}{(x_i - \bar{x})^2} \quad (4)$$

$$\hat{\beta}_0 = \bar{Y} - \beta_1 \bar{X} \quad (5)$$

Where $\hat{\beta}_1$ & $\hat{\beta}_0$ are estimated values of the regression parameters β_1 & β_0 and $\bar{X}$, $\bar{Y}$ are mean values of X & Y.

A. Data Source

Sample Data is taken from Advertising Ratios & Budgets provided in annual report by Schonfeld & Associates, Inc [6]. which covers over 2,400 companies and 320 industries with information on fiscal 2018 and 2019 advertising& revenue spending.

For OLS Analysis total 52 sample companies data is taken from 12 different industries.

Table - 1 shows the sample percentage revenue growth & percentage advertisement growth for Electromedical & Electrotheraputic Appartus.

Growth is taken from 2018 to 2019			Adv Grw & Rev Grw are in %
Sr.No	Company	Ad Grw	Rev Grw
1	Adm Tronics Unlimited, Inc.	-3.88	-19.58
2	Axogen, Inc.	-85.78	27.13
3	Biolife Solutions Inc	43.33	38.64
4	Cutera Inc	0	11.67
5	Digirad Corp	0	9.6
6	Edap Tms Sa	-2.7	8.51
7	Electromed, Inc.	21.52	10.57
8	Escalon Medical Corp	22.22	-15.57
9	Fonar Corp	-11.37	6.96
10	Iridex Corp	-40	1.99
11	Masimo Corp	-21.79	9.27
12	Medifirst Solutions, Inc.	-92.13	-2.73

Table – 1:Data Source[6]: June 2020 Sample data of Advertising Ratios & Budgets from Schonfeld-Associates-Inc-v417 of Market Research.com

We will develop an simple regression model to understand and predictpercentage sales revenue growth on the percentage advertisement growth.

B. Creating Feature Set(X) and Outcome Variable(Y) Using Python

The OLS model takes two parameters Y and X.In our example percentage advertisement growth will be X and percentage sales revenue growth will be Y.We will split data set into two sets, training & validation set. Trainng set will be used to train algorithm to predict output. Validation set will be used to test accuracy & efficiency.

C. Python for Building Regression Model

Python language is used as tool for building regression model for sales prediction. The statsmodel library is used in Python for building statistical models. OLS(Oridnary least square) API available in statsmodel.api is used for estimation of parameters for simple linear rgression model.

D. Splitting the Dataset into Training and Validation Set

The data is divided into two subsets training data set and validation data set. The proportion of training dataset is usually between 70% and 80% of the data and the remaining data is used for validation data. We have taken train_size = 0.8which implies that 80% of the data will be used for training the model and remaining 20% will be used for validating the model. The records that are selected for training and test set are randomly sampled using python functions which returns four variables as shown below.

train_X = feature values of the training set

train_Y = response values of the training set

test_X = feature values of the test set

test_Y = response values of the test set

E. Finding estimated parameters

Regression parameters $\hat{\beta}_1 \& \hat{\beta}_0$ are estimated from equations (4) & (5) using Python functions as tool.

F. Fitting the Model

Linear regression calculates an equation that minimizes the distance between the fitted line and all of the observed data points. Technically, ordinary least squares (OLS) regression minimizes the sum of the squared residuals. In general, a model fits the data well if the differences between the observed values and the model's predicted values are small and unbiased.

G. Co-efficient of Determination (R-Squared / R^2)

R-squared is a statistical measure of how close the data are to the fitted regression line. It is defined as

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}} \qquad (6)$$

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2} \qquad (7)$$

SSR = The *sum squared regression (SSR)* is the sum of the square residuals $(y_i - \hat{y}_i)^2$. Residual is the difference between observed value y_i & estimated value $\hat{y}_i$ as shown below in Fig - 1.

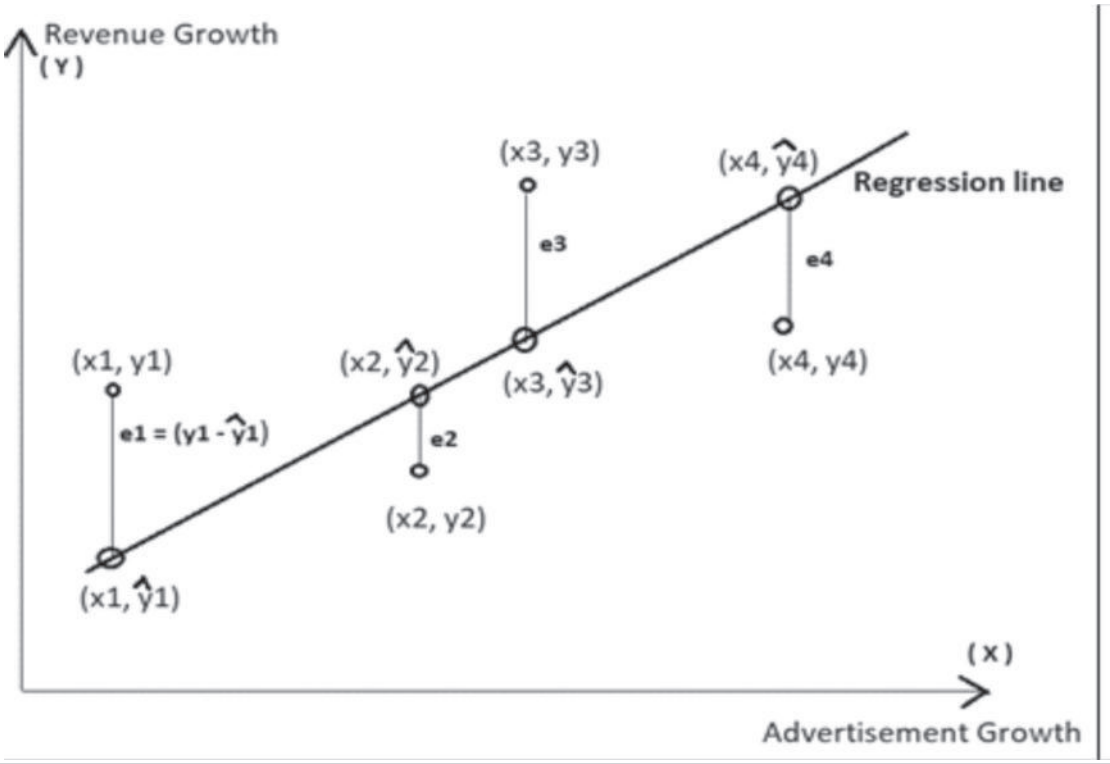


Fig – 1: Residuals as function of Actual value & estimated value

$$SSR = \sum(y_i - \hat{y}_i)^2 = e_1^2 + e_2^2 + e_3^2 + e_4^2 \quad (8)$$

= square sum of variations w.r.t to estimated value

SST = *total sum of squares* is the sum of the distance the data is away from the mean (central tendency) all squared as shown below in Fig - 2.

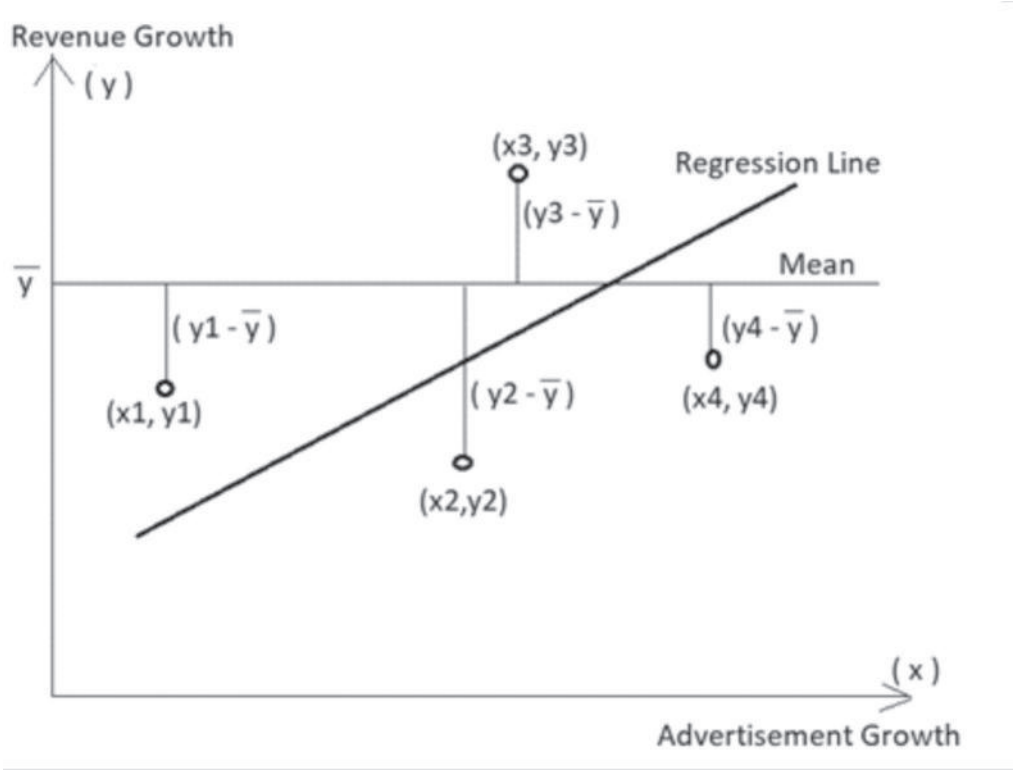


Fig – 2: Residuals as function of Actual value & Mean value

$$SST = \sum(y_i - \bar{y})^2 = (y1 - \bar{y})^2 + (y2 - \bar{y})^2 + (y3 - \bar{y})^2 + (y4 - \bar{y})^2 \quad (9)$$

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}}$$

$$R^2 = \frac{SST - SSR}{SST}$$

$$R^2 = \frac{\sum(y_i - \bar{y})^2 - \sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (10)$$

The above equation indicates that R2 is directly proportional to difference between the square sum of variations in y w.r.t mean and square sum of variations in y w.r.t estimated value.

Not good fit:

Smaller R2 value indicates that SSR value is large and close to SST which indicates that variation in y w.r.t estimated value is large & close to variation in y w.r.t mean, which is not good fit.

Good fit:

Large R2 value indicates that SSR value is very small (actual values of y are close to estimated values of y) and not close to SST, which indicates that variation in y w.r.t estimated value is not close to variation in y w.r.t mean, which is a good fit.

Results & Model Diagnostics

A. Estimated Model

Using python as tool parameters of regression model are calculated as shown below.

Using 80% training data set

- ✓ Constant $\hat{\beta_0}$ = 6.101
- ✓ Regression coefficient $\hat{\beta_1}$ = 0.160

The estimated model can be written as

$Y_i = \beta_0 + \beta_1 X_i(11)$

$$\text{Rev Grw}(\%) = 6.101 + 0.160 * (\text{Ad Grw}(\%))$$

B.Interpretation of Estimated Model

Model estimates that 1% Ad Growth will increase Revenue by 0.160 %. For example, if the sales revenue was 2 Million in year 2018 then according to our model sales revenue in year 2019 will increase by 0.0032 million i.e. estimated sales revenue can be 2.0032 millions that is rise of 3200/- in revenue.

C.Model Diagnostics (Validation)

Before using regression model in practical applications, it should be validated & tested for goodness of fit. We will be using Co-efficient of determination (R-squared) method to determine goodness of fit. Using python as a tool following value of R2 is calculated

R2 = 0.208

According to Cohen – 1992 [9] r-square value 0.12 (12%) or below indicate low, between 0.13 (13%) to 0.25 (25%) values indicate medium& 0.26 (26%) or above values indicate high. Our model explains 20.8% of the variance in the validation set, so it is reasonably good fit.

Conclusion

The simple linear regression model using ordinary least square (OLS) method shows functional relationship between the outcome variable (Sales revenue growth in %) and the feature (advertisement growth in %). The model validation is investigated using R2 technique to ensure goodness of fit.while an R-square as low as 10% is generally accepted for studies in the field of arts, humanities and social sciences because human behavior cannot be accurately predicted, therefore, a low R-square is often not a problem in studies in the arts, humanities and social science field. There are various other control parameters which affects the value of R-square. Therefore, in order to extend scope of this research various social science characteristics like age, gender, motivation towards product and festive season should be included as control variables in analysis.

References:

[1] Core Python Programming by Dr.R.Nageswara Rao second edition dreamtechPress.

[2] Machine Learning Using Python by Manaranjan Pradhan & U Dinesh Kumar first reprint edition Wiley publications

[3] Sales prediction types available online at URL: https://www.economicsdiscussion.net/sales/sales-forecasting-methods/32270

[4] Advantages of Linear regression model available online at URL: https://iq.opengenus.org/advantages-and-disadvantages-of-linear-regression/

[5] Will Koehrsen Article how to setup your-machine learning problem can be found at following URL:https://towardsdatascience.com/prediction-engineering-how-to-set-up-your-machine-learning-problem-b3b8f622683b

[6] Data source of Ratios & Budgets can be found at following URL:

https://www.marketresearch.com/Schonfeld-Associates-Inc-v417/Advertising-Ratios- Budgets-13373044/

[7] https://www.keboola.com/blog/linear-regression-machine-learning.

[8] https://internal.ncl.ac.uk/ask/numeracy-maths-statistics/statistics/regression-and-correlation/coefficient of-determination-r-squared.html#:~:text=R2%3D1%E2%88%92sum%20squared,from%20the%20mean%20all%20squared.

[9] Cohen's Conventions for Small, Medium,and Large R2 values can be found on following URL http://core.ecu.edu/psyc/wuenschk/docs30/EffectSizeConventions.pdf

[10] Small is beautiful. The use and interpretation of R2 in social Research by Ferenc Moksony Pages 6 & 7

Certificates

The Institution of Engineers (India)

NAVI MUMBAI LOCAL CENTRE

and

Agnel Charities'

FR. C. RODRIGUES INSTITUTE OF TECHNOLOGY, VASHI

Jointly present

IEI NMLC-FCRIT EXCELLENCE AWARDS - 2023

Certificate of Participation

This is to certify that *Mr. Avinash Dangwani* of **Pacific Academy of Higher Education and Research University, Udaipur** has participated in ***IEI NMLC-FCRIT EXCELLENCE AWARDS*** ceremony held on September 15th, 2023 and the nomination is found to be worthy for consideration under **Research Excellence in Students (PhD) Regional Category.**

DR. NILAJ DESHMUKH
CHAIR, IEI NMLC-FCRIT
EXCELLENCE AWARDS-2023

DR. S. M. KHOT
PRINCIPAL,
FCRIT, VASHI

ER. SANJAY R. BAGUL
SECRETARY,
IEI, NMLC

DR. K. M. GODBOLE
CHAIRMAN,
IEI, NMLC

Elite

NPTEL Online Certification

(Funded by the MoE, Govt. of India)

TOP TOPPER 5%

This certificate is awarded to

AVINASH DANGWANI

for successfully completing the course

Python for Data Science

with a consolidated score of **78** %

Online Assignments	21.67/25	Proctored Exam	56.67/75

Total number of candidates certified in this course: **7596**

Devendra Jalihal

Prof. Devendra Jalihal
Chairperson,
Centre for Outreach and Digital Education, IITM

Jul-Aug 2023

(4 week course)

Prof. Andrew Thangaraj
NPTEL, Coordinator
IIT Madras

Indian Institute of Technology Madras

FREE ONLINE EDUCATION swayam शिक्षित भारत, उन्नत भारत

Roll No: NPTEL23CS99S35002545

To verify the certificate

No. of credits recommended: 1 or 2

Elite

NPTEL Online Certification

(Funded by the MoE, Govt. of India)

This certificate is awarded to

AVINASH DANGWANI

for successfully completing the course

Data Analytics with Python

with a consolidated score of **82** %

Online Assignments	24.69/25	Proctored Exam	57.45/75

Total number of candidates certified in this course: **7860**

Prof. Sanjeev Manhas
Coordinator, Continuing Education Centre
IIT Roorkee

Jan-Apr 2023

(12 week course)

Priti Maheshwari

Prof. Priti Maheshwari
NPTEL Coordinator
IIT Roorkee

Indian Institute of Technology Roorkee

FREE ONLINE EDUCATION swayam शिक्षित भारत, उन्नत भारत

Roll No: NPTEL23CS08S54221311

To validate the certificate

No. of credits recommended: 3 or 4

DEPARTMENT OF PG STUDIES AND DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

(Pacific Academy of Higher Education and Research University, Udaipur)

INTERNATIONAL VIRTUAL CONFERENCE
on
EMERGING ERA OF APPLICATIONS OF COMPUTER

Certificate of Merit

This is to certify that *Mr. Avinash Sunder Dangwani*

of *Pacific Academy of Higher Education & Research University*

has been awarded as Young Scientist for Best Oral Presentation in this International Virtual Conference on Emerging Era of Applications of Computer on 15th-16th January 2022.

Prof. Hemant Kothari	**Dr. Dilendra Hiran**	**Dr. Tanveer Ahmed Kazi**	**Dr. Prashant Sharma**
Conference Chairman	Conference Co-Chairman	Conference Director	Conference Organizing Secretary
PAHER University, Udaipur	PAHER University, Udaipur	PAHER University, Udaipur	PAHER University, Udaipur

PACIFIC ACADEMY OF HIGHER EDUCATION AND RESEARCH UNIVERSITY, UDAIPUR

IQAC Cell

in association with

PACIFIC BUSINESS SCHOOL, UDAIPUR

Faculty Development Program

January 2-7, 2021

Certificate of Participation

This is certified that

Dr./Ms./Mr. **Avinash Dangwani, Associate Professor (Selection Grade)**

of **VES Polytechnic**

has successfully participated in

Faculty Development Program on 'Research and Publication Ethics'

Prof. K.K. Dave
President, PAHER University
(FDP Chair)

Prof. Hemant Kothari
Dean, P.G. Studies, PAHER University
(Convener FDP)

Prof. Mahima Birla
Dean, Faculty of Management
(Director FDP)

Prof. Dipin Mathur
Director, Pacific Business School
(Co-convener FDP)

EDUXLABS
edu's Xpert (Esoir Solution)

CERTIFICATE OF COMPLETION

THIS IS TO CERTIFY THAT DR./PROF./MR./MS.

Avinash Dangwani

has successfully completed Live Online Instructor-led
10-Days (35 Hours) Faculty Development Program / Short Training on

Applied Deep Learning for Medical Data Analysis

[MRI, CT-SCAN, ECG, & COVID-19 Chest X-RAY]

From 27th October to 9th November, 2020

Organized by

EduxLabs (Esoir Business Solution LLP)

ISSUED ON:
10TH NOVEMBER 2020

Esoir Business Solution LLP
MD.NAFISH Director
EDUXLABS DIRECTOR
(ESOIR BUSINESS SOLUTION)