ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 1, January 2024

Choosing the Optimal Model When Forecasting the Number of Patients with Malignant Tumor Disease

Nurmukhammad Alimkulov, Mirzaakbar Hudayberdiev

PhD student, Andijan State University, Andijan, Uzbekistan

Professor of Department of Information Technology Software, Tashkent University of Information Technologies, Tashkent, Uzbekistan

ABSTRACT: The acquisition of predictive cancer data holds paramount importance for comprehending the multifaceted dimensions of cancer incidence. This imperative extends beyond mere cancer burden estimation, encompassing critical aspects such as healthcare planning and resource allocation. In this context, the Global Cancer Observatory (Globocan) emerges as a pivotal program, monitoring cancer incidence worldwide. Discrepancies between Globocan's projections and official statistics underscore the significance of accurate prediction. Predictive models play a pivotal role in healthcare, necessitating meticulous development and ethical considerations. This study explores the selection of an apt time series model for forecasting new cancer cases, with a focus on the Long Short-Term Memory (LSTM) model's performance. The LSTM model for sequential data is an effective tool in cancer prediction.

I. INTRODUCTION

There exists a compelling imperative for the acquisition of predictive cancer data, motivated by multifaceted considerations. Specifically, the necessity for comprehensive insights into cancer incidence manifests to elucidate the intricate cancer burden dynamics, involving the systematic monitoring of cancer prevalence across temporal trajectories and diverse demographic cohorts. This knowledge reservoir serves as a pivotal instrument for the identification of regions characterized by elevated cancer incidence rates, thereby facilitating the evaluation of the efficacy of interventions in cancer prevention and control programs. Significantly, within the domain of healthcare planning, the discipline of Cancer Predictions assumes a pivotal role. It empowers healthcare entities with the foresight essential for provisioning for the impending requirements of cancer patients, encompassing the critical facets of staffing adequacy with specialized oncologists, nurses, and allied healthcare professionals, as well as the requisite provisioning of essential medical equipment and supplies [1], [2].

Prudent consideration of data confidentiality is of paramount import in the realm of cancer prediction. Predictive models necessitate meticulous development, cognizant of multifarious factors encompassing data quality, model precision, and susceptibility to error. Moreover, the requisite urgency for cancer prediction is inherently contingent upon contextual nuances, varying across geographic regions, cancer typologies, and the availability of healthcare resources. These distinct contextual variances might necessitate tailored emphases, pivoting between preventive paradigms in certain contexts and proactive early detection and treatment strategies in others [3], [4].

In the pursuit of global cancer surveillance and intelligence dissemination, the Global Cancer Observatory (Globocan) has emerged as a seminal program. Conceived and advanced by the International Agency for Research on Cancer (IARC), an integral component of the esteemed World Health Organization (WHO), Globocan is aptly abbreviated as the "Global Cancer Observatory." Its overarching mission resides in the systematic monitoring of cancer incidence, mortality, and morbidity on a global scale. In alignment with this mandate, Globocan is committed to furnishing comprehensive information pertaining to the prevalence and dynamics of cancer across diverse global settings [5].

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 1, January 2024

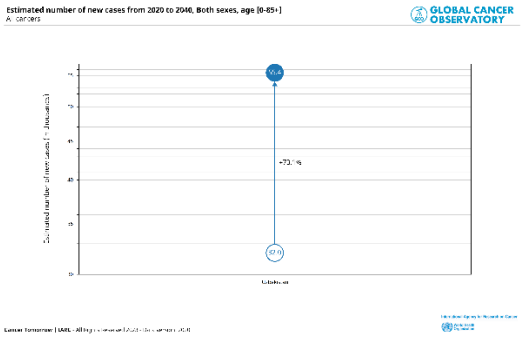


Fig.1. Illustrates Globocan's projected metrics concerning the prospective identification of new patients within the Republic of Uzbekistan spanning the years from 2020 to 2040.

As per Globocan's estimations, the anticipated count of newly identified cases within the Republic of Uzbekistan for the year 2020 stood at a minimum threshold of 32,000 cases. In stark contrast, official statistics report a notably lower figure of 21,976 cases. This substantial disparity underscores a noteworthy increment in the count of cases, signifying that a considerable portion of cases may have gone unaccounted for or unreported.

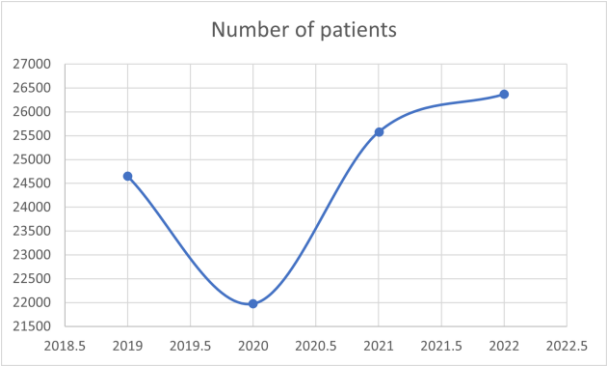


Fig.2. The number of newly diagnosed patients in the Republic of Uzbekistan in 2019-2022.

II. LITERATURA REVIEW

Precision in forecasting cancer incidence emerges as a pivotal imperative in confronting the multifaceted challenges inherent to this intricate and debilitating malady. As previously expounded, these prognostications wield profound significance, transcending mere estimation of cancer burden to encompass pivotal domains such as healthcare strategization, resource allocation, and meticulous health provisioning.

Cancer, characterized by its polygenic etiology, mandates meticulous strategization for the optimal delivery of patient care. The Significance of Precision Predictions materializes across several pivotal dimensions:

- Resource Allocation: Within the healthcare milieu, the allocation of resources stands as a salient concern. Precision forecasts serve as the bedrock for ensuring that healthcare facilities are staffed judiciously with the requisite cadre of specialized oncologists, nurses, and ancillary support staff. Furthermore, anticipatory insights into expected patient influxes engender proactive measures, ensuring the timely provisioning of essential medical equipment and requisite supplies. This proactive stance obviates logistical bottlenecks and augments the efficacy of healthcare delivery.
- Progress Monitoring: Predictive models assume a pivotal role in the ongoing evaluation of the efficacy of cancer prevention and control programs. A comparative analysis between forecasted and observed cancer incidence rates empowers policymakers to scrutinize the impact of their interventions, enabling data-driven adaptations and refinements as necessitated by evolving circumstances.
- Research and Development: Precision predictions furnish invaluable insights for cancer researchers, facilitating

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 1, January 2024

the identification of emerging trends and potential risk factors. These prognostications, underpinned by meticulous data analysis, form the cornerstone for further investigative endeavors. The outcomes hold the promise of fostering advancements in the realms of cancer prevention and treatment.

- Ethical Contemplations: While predictive models hold the potential for immense benefits, they concurrently engender ethical deliberations surrounding data privacy and confidentiality. Striking the delicate equilibrium between predictive precision and the imperative of safeguarding sensitive patient information constitutes a perennial challenge in the realm of cancer prognostication. Ethical considerations demand continuous attention to ensure the responsible utilization of predictive technologies.

Precision in cancer incidence predictions assumes the mantle of a linchpin, guiding strategic healthcare planning, prudent resource management, and data-informed policymaking. It serves as a beacon illuminating the path towards enhanced patient care, research breakthroughs, and ethically sound practices in the ever-evolving landscape of cancer control and prevention [6], [7].

III. PROPOSED WORKFLOW

The judicious selection of an apt time series model stands as a pivotal stride in ensuring the precision of cancer prognostication. This choice is contingent upon a multifaceted spectrum of variables, encompassing the nuanced objectives of the forecast, the caliber and volume of the accessible data, and its sufficiency to meet the requisite modeling standards.

This scholarly inquiry embarks upon a meticulous expedition, elucidating the labyrinthine process of singling out an apt time series model tailored for the prognostication of the influx of new cancer patients. In this rigorous examination, we meticulously scrutinized a gamut of modeling methodologies, encompassing classical statistical paradigms and cutting-edge machine learning models of heightened sophistication. Our unswerving focus has been on the attainment of the acme of predictive precision, a mission that augments the arsenal of tools available for judicious healthcare strategization and decision-making.

Subsequent to this prologue, our exposition proceeds to unveil the fruits of our analytical labor. We expound upon the outcomes of our investigation, unraveling the performance metrics of the Long Short-Term Memory (LSTM) model, a neural network architecture custom-crafted for sequential data. This model, bearing the potential to discern intricate temporal dependencies, beckons toward its possible deployment as a formidable asset in the annals of cancer prediction, a prospect we dissect and explore in the forthcoming sections.

IV. MODEL ARCHITECTURE

In the pursuit of crafting a precise and resource-efficient time series model for the prognostication of the burgeoning tide of novel cancer cases, a comprehensive survey of modeling options was undertaken. Among the array of models subject to contemplation, the Long Short-Term Memory (LSTM) model emerged as a compelling selection, distinguished by its distinctive neural architecture and formidable computational prowess [8].

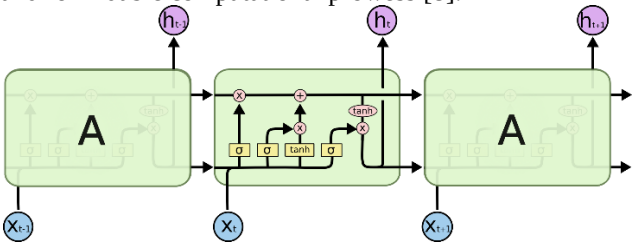
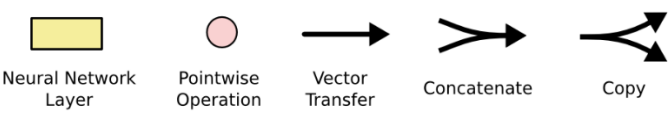


Fig.3. Depicts the intrinsic architecture of the recurrent module within the LSTM model, characterized by the intricate interplay of four interconnected layers.

here,



ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 1, January 2024

Long Short-Term Memory (LSTM) emerges as a specific category within the realm of recurrent neural networks (RNNs), distinguished for its aptitude in processing sequences of data, rendering it an exemplary choice for the modeling of time-series data, such as cancer incidence. The primary distinction that confers LSTM its superiority over conventional time series models lies in its unique capacity to apprehend intricate patterns and interrelationships that span prolonged temporal horizons. This capacity significantly improves the predicament of the vanishing gradient, a persistent challenge encountered in standard RNNs.

Progressing within this exposition, we delve into a comprehensive scrutiny of the LSTM model's performance, particularly in the context of its prognostic abilities concerning the number of newly diagnosed cancer patients. Our overarching objective remains twofold: firstly, to substantiate its efficacy as a predictive instrument and secondly, to cast illumination upon its potential utility within the expansive domain of cancer prediction [9].

The foundational architecture of this predictive model encapsulates an initial input layer, artfully crafted to ingest and process sequential data. This conduit, in turn, channels the sequential data to a follow-on sequential LSTM layer, characterized by an assembly of 50 units. The LSTM layer's intrinsic design intricately intertwines with the data, facilitating the capture of intricate temporal relationships embedded within the input sequence. Ultimately, an output layer assumes the mantle of generating predictions predicated upon the information gleaned from the LSTM layer's output. This architectural construct bestows upon the model a heightened competence in the discernment and analysis of temporal dynamics inherent to time series data, rendering it eminently suited for the task of prognosticating the incidence of new cancer cases.

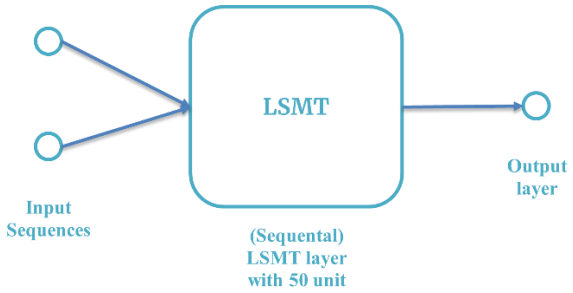


Fig.4. A model developed for prediction.

V. RESULTS AND DISCUSSION

The dataset under consideration for this study encompasses the tally of newly diagnosed cancer patients within the Andijan region, spanning from the inception of January 2023 through to the culmination of August 2023. This meticulously curated dataset serves as the foundational bedrock upon which our predictive modeling endeavors are erected. Employing the bespoke prediction model meticulously developed for this study, we have undertaken the task of generating forecasts pertaining to the anticipated count of newly diagnosed cancer patients within the confines of the Andijan region over the ensuing two-month temporal horizon. This concerted effort augments our comprehension of the intricate temporal dynamics characterizing the landscape of cancer incidence within this specific geographic demesne.

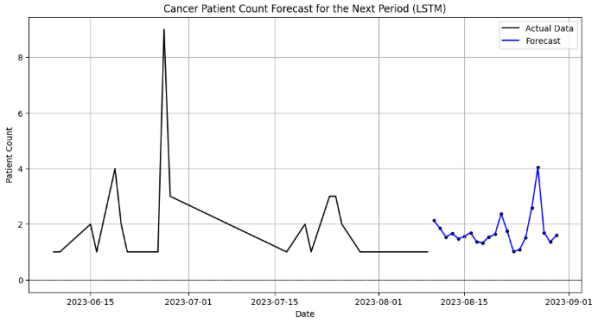


Fig.5. Next two months prediction using LSTM.

The discussion delves into the critical importance of precise cancer incidence predictions in healthcare planning and

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 11, Issue 1, January 2024

decision-making. It highlights the multifaceted nature of cancer as a disease, necessitating tailored approaches for patient care. Resource allocation and progress monitoring are underscored as key domains influenced by predictive accuracy. Research and development benefit from precise predictions, enabling the identification of trends and risk factors. Ethical considerations surrounding data privacy are acknowledged as a constant challenge.

VI. CONCLUSION

In conclusion, this study underscores the pivotal role of accurate cancer predictions in healthcare and beyond. It sheds light on the importance of Globocan's global surveillance efforts and highlights discrepancies in reported cancer cases. The study also emphasizes the ethical considerations surrounding predictive models. Finally, it explores the potential of the LSTM model for cancer prediction, offering a promising avenue for further research and applications in the field.

VII. ACKNOWLEDGEMENT

We extend our profound appreciation to the Innovative Development Agency under the aegis of the Ministry of Higher Education, Science, and Innovation for their unwavering support in facilitating this research endeavor, operating within the ambit of practical project No. FZ-202010191.

REFERENCES

[1] Haitham Elwahsh, Medhat A. Tawfeek, A.A. Abd El-Aziz, Mahmood A. Mahmood, Maazen Alsabaan, Engy El-shafeiy, "A new approach for cancer prediction based on deep neural learning," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 6, p. 101565, 2023.

[2] Varsha Nemade, Vishal Fegade, "Machine Learning Techniques for Breast Cancer Prediction," *Procedia Computer Science*, vol. 218, pp. 1314-1320, 2023.

[3] Xiao-Jing Luo, Qi Zhao, Jia Liu, Jia-Bo Zheng, Miao-Zhen Qiu, Huai-Qiang Ju, Rui-Hua Xu, "Novel Genetic and Epigenetic Biomarkers of Prognostic and Predictive Significance in Stage II/III Colorectal Cancer," *Molecular Therapy*, vol. 29, no. 2, pp. 587-596, 2021.

[4] Amanda F. Petrik, Eric S. Johnson, Rajasekhara Mummadi, Matthew Slaughter, Gloria D. Coronado, Sunny C. Lin, Lucy Savitz, Neal Wallace, "The use of individual and multilevel data in the development of a risk prediction model to predict patients' likelihood of completing colorectal cancer screening," *Preventive Medicine Reports*, vol. 36, p. 102366, 2023.

[5] Deependra Singh, Jerome Vignat, Valentina Lorenzoni, Marzieh Eslahi, Ophira Ginsburg, Beatrice Lauby-Secretan, Marc Arbyn, Partha Basu, Freddie Bray, Salvatore Vaccarella, "Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative," *The Lancet Global Health*, vol. 11, no. 2, pp. e197-e206, 2023.

[6] Cong Wang, Mengyan Zhang, Jiyun Zhao, Bin Li, Xingjun Xiao, Yan Zhang, "The prediction of drug sensitivity by multi-omics fusion reveals the heterogeneity of drug response in pan-cancer," *Computers in Biology and Medicine*, vol. 163, p. 107220, 2023.

[7] Mohamed Omar, Wikum Dinalankara, Lotte Mulder, Tendai Coady, Claudio Zanettini, Eddie Luidy Imada, Laurent Younes, Donald Geman, Luigi Marchionni, "Using biological constraints to improve prediction in precision oncology," *iScience*, vol. 26, no. 3, p. 106108, 2023.

[8] Olah, Christopher, "Understanding LSTM Networks," 27 08 2015. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/?source=post_page-----37e2f46f1714----------.

[9] Rinku Supakar, Parthasarathi Satvaya, Prasun Chakrabarti, "A deep learning based model using RNN-LSTM for the Detection of Schizophrenia from EEG data," *Computers in Biology and Medicine*, vol. 151, no. A, p. 106225, 2022.