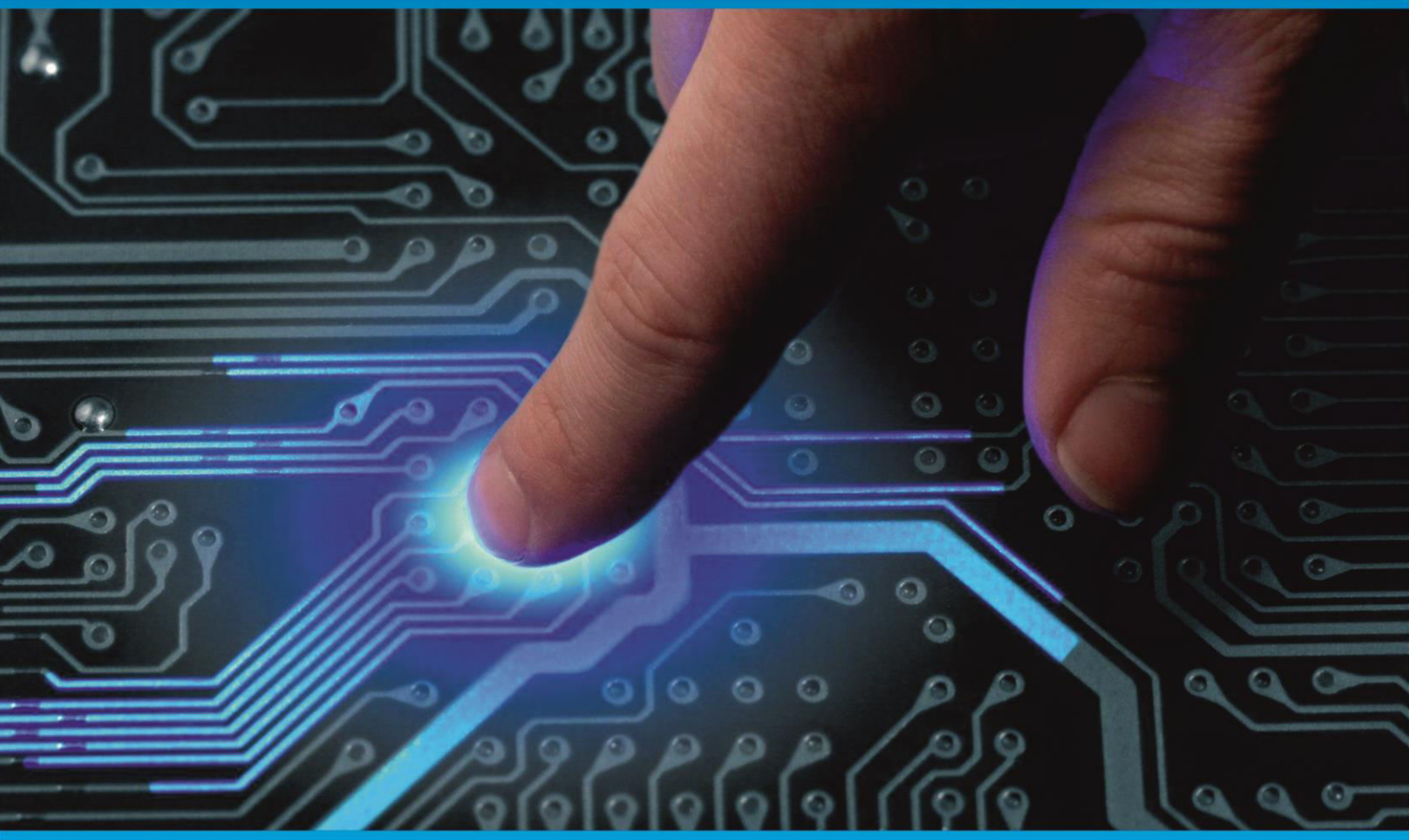
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 3, March 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

Secure Screening Test for the Patent of New Drug Discovery using Cloud Computing

Eswara Moorthy P[1], Lokesh R[2], A.Anbumani M.E[3]

UG Student, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai, Tamilnadu, India[1,2]

Assistant Professor, Department of Computer Science and Engineering, Velammal Institute of Technology, Chennai, Tamilnadu, India[3]

ABSTRACT: In this project, Secure screening test is designed to allow the cloud to securely use multiple drug formulas to check whether the new drug compounds is existing or not in the Existing Database. We use multiple existing drug formulas to train Support Vector Machine (SVM) provided by the analytical model provider. In our approach, we design secure computation protocols to allow the cloud server to perform commonly used integer and fraction computations. To securely train the SVM, we design a secure SVM parameter selection protocol to select two SVM parameters and construct a secure sequential minimal optimization protocol to privately refresh both selected SVM parameters. The trained SVM classifier can be used to determine whether a drug chemical compound is active or not in a privacy-preserving way. Here Naive Bayes algorithm is used for quick prediction. SVM and NB algorithm gives trained data and accuracy, but when there is a large dataset its accuracy may vary. So NB is used for quick prediction. Lastly, we prove that the proposed POD achieves the goal of SVM training and chemical compound classification without privacy leakage to unauthorized parties.

KEYWORDS: Secure screening test, new drug compounds, Support Vector Machine, quick prediction

I. INTRODUCTION

The aim of the project is to allow the cloud to securely use multiple drug formulas for screening test from the drug providers with the trained datasets. In this project, Secure screening test is designed to allow the cloud to securely use multiple drug formulas to check whether the new drug compound is exist or not in the Existing Database. We use multiple existing drug formulas to train Support Vector Machine (SVM) provided by the analytical model provider. In our approach, we design secure computation protocols to allow the cloud server to perform commonly used integer and fraction computations. To securely train the SVM, we design a secure SVM parameter selection protocol to select two SVM parameters and construct a secure sequential minimal optimization protocol to privately refresh bothselected SVM parameters. The trained SVM classifier can be used to determine whether a drug chemical compound is active or not in a privacy-preserving way. Here Naive Bayes algorithm is used for quick prediction. Lastly, we prove that the proposed POD achieves the goal of SVM training and chemical compound classification without privacy leakage to unauthorized parties.

II. LITERATURE SURVEY

A literature review is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources and do not report new or original experimental work.

Title : Deep Learning in Drug Discovery and Medicine Scratching the Surface.
Author Name : Dibyendu Dana, Satishkumar V.Gadhiya , Luce G. St..Surin , David Li , Farha Naaz ,Quaisar Ali
Year of Publication: 2018

Abstract:

The practice of medicine is ever evolving. Diagnosing disease, which is often the first step in a cure, has seen a sea change from the discerning hands of the neighborhood physician to the use of sophisticated machines to use of

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

information gleaned from biomarkers obtained by the most minimally invasive of means. The last 100 or so years have borne witness to the enormous success story of allopathy, a practicethat found favor over earlier practices of medical purgatory and homeopathy. Nevertheless, failures of this approach coupled with the omicsand bioinformatics revolution spurred precision medicine, a platform wherein the molecular profile of an individual patient drives the selection of therapy. Indeed, precision medicine-based therapies that first found their place in oncology are rapidly finding uses in autoimmune, renal and other diseases. More recently a new renaissance that is shaping everyday life is making its way into healthcare. Drugdiscovery and medicine that started with Ayurveda in India are now benefiting from an altogether different artificial intelligence (AI)—one which is automating the invention of new chemical entities and the mining of large databases in health-privacy-protected vaults. Indeed, disciplines as diverse as language, neurophysiology, chemistry, toxicology, biostatistics, medicine and computing have come together to harness algorithms based on transfer learning and recurrent neural networks to design novel drug candidates, a priori inform on their safety, metabolism and clearance, and engineer their delivery but only on demand, all the while cataloging and comparing omics signatures across traditionally classified diseases to enable basket treatment strategies. This review highlights inroads made and being made in directed-drugdesign and molecular therapy.

Title : A Novel Neutrosophic Weighted Extreme Learning Machine for Imbalanced Data Set
Author Name : Yaman Akbulut 1 ID , Abdulkadir ¸Sengür 1,* ID , Yanhui Guo 2 and Florentin Smarandache
Year of Publication : 2017

Abstract:

Extreme learning machine (ELM) is known as a kind of single-hidden layer feedforward network (SLFN), and has obtained considerable attention within the machine learning community and achieved various real-world applications. It has advantages such as good generalization performance, fast learning speed, and low computational cost. However, the ELM might have problems in the classification of imbalanced data sets. In this paper, we present a novel weighted ELM scheme based on neutrosophic set theory, denoted as neutrosophic weighted extreme learning machine (NWELM), in which neutrosophic c-means (NCM) clustering algorithm is used for the approximation of the output weights of the ELM. We also investigate and compare NWELM with several weighted algorithms. The proposed method demonstrates advantages to compare with the previous studies on benchmarks.

Title : Drug Design and Discovery Principles and Applications
Author Name : Shu-Feng Zhou , Wei-Zhu Zhong
Year of Publication : 2017

Abstract:

Drug discovery is the process through which potential new therapeutic entities are identified, using a combination of computational, experimental, translational, and clinical models Despite advances in biotechnology and understanding of biological systems, drug discovery is still a lengthy, costly, difficult, and inefficient process with a high attrition rate of new therapeutic discovery. Drug design is the inventive process of finding new medications based on the knowledge of a biological target. In the most basic sense, drug design involves the design of molecules that are complementary in shape and charge to the molecular target with which they interact and bind. Drug design frequently but not necessarily relies on computer modeling techniques and bioinformatics approaches in the big data era. In addition to small molecules, biopharmaceuticals and especially therapeutic antibodies are an increasingly important class of drugs and computational methods for improving the affinity, selectivity, and stability of these protein-based therapeutics have also gained great advances [3]. Drug development and discovery includes preclinical research on cell-based and animal models and clinical trials on humans,and finally move forward to the step of obtaining regulatory approval in order to market the drug. Modern drug discovery involves the identification of screening hits, medicinal chemistry and optimization of those hits to increase the affinity, selectivity (to reduce the potential of side effects), efficacy/potency, metabolic stability (to increase the half-life), and oral bioavailability. Once a compound that fulfills all of these requirements has been identified, it will begin the process of drug development prior to clinical trials.

III. PROPOSED METHODOLOGY

we propose a **p**rivacy preserving **O**utsourced Support Vector Machine Design for Secure **D**rug discovery in the cloud environment. Unlike existing drug discovery frameworks, our POD seeks to achieve it efficiently and securely. We are not using three real time datasets to check the efficiency of potential new drug component. Instead of using existing

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

datasets we are using SVM and another one data mining algorithm Naive Bayes(NB). These two algorithms are used to train the uploaded drug dataset (CSV file). In final we will get trained dataand accuracy for that uploaded dataset. Drug tester will check that new drug component. Drug tester doesn't know the contents of that file;they will get the trained data only. Then they let us know the file was active or not. And finally, admin will approve the drug component.

Advantage

- We minimize the risk of unauthorized disclosure during the SVM and NB training.
- Multiple pharmaceutical corporations won't reveal the drug components in detail.

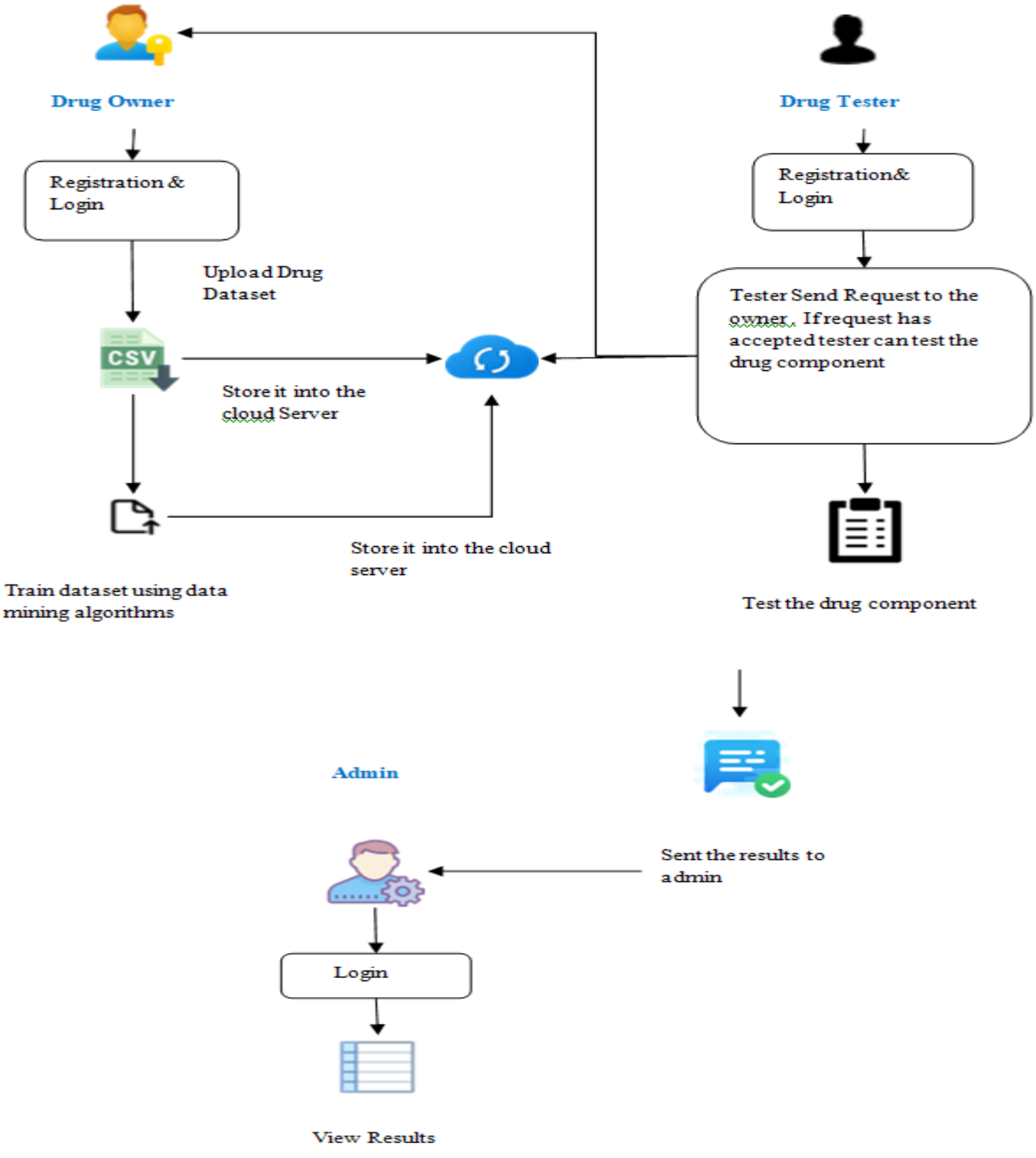


Fig.1 Architecture Diagram

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

Drug Owner & Tester Registration

The drug owner and drug tester will register their details. Both of them should register their personal details. Those details will be stored into the database.

Drug Component Uploading

The drug owner will upload the data set. That data set contains the chemical components and we have to mention the type of class (Class A, Class B). While uploading the file drug owner need to enter the unique drug name and id, we will read the content and store into the database and store that .csv file in cloud.

Train dataset

While uploading itself the algorithm will train the uploaded data using python. For this part we will use two algorithms SVM and Naive Bayes. The trained data and accuracy will be sent to the owner from python server.

Drug Testing

The drug tester will test the uploaded drug components. The Tester will test that particular drug component is still active or not in the database and send the results to the Admin.

IV. EXPERIMENTAL RESULTS

The following figures shows process of secure screening test for the patent of new drug discovery using cloud computing

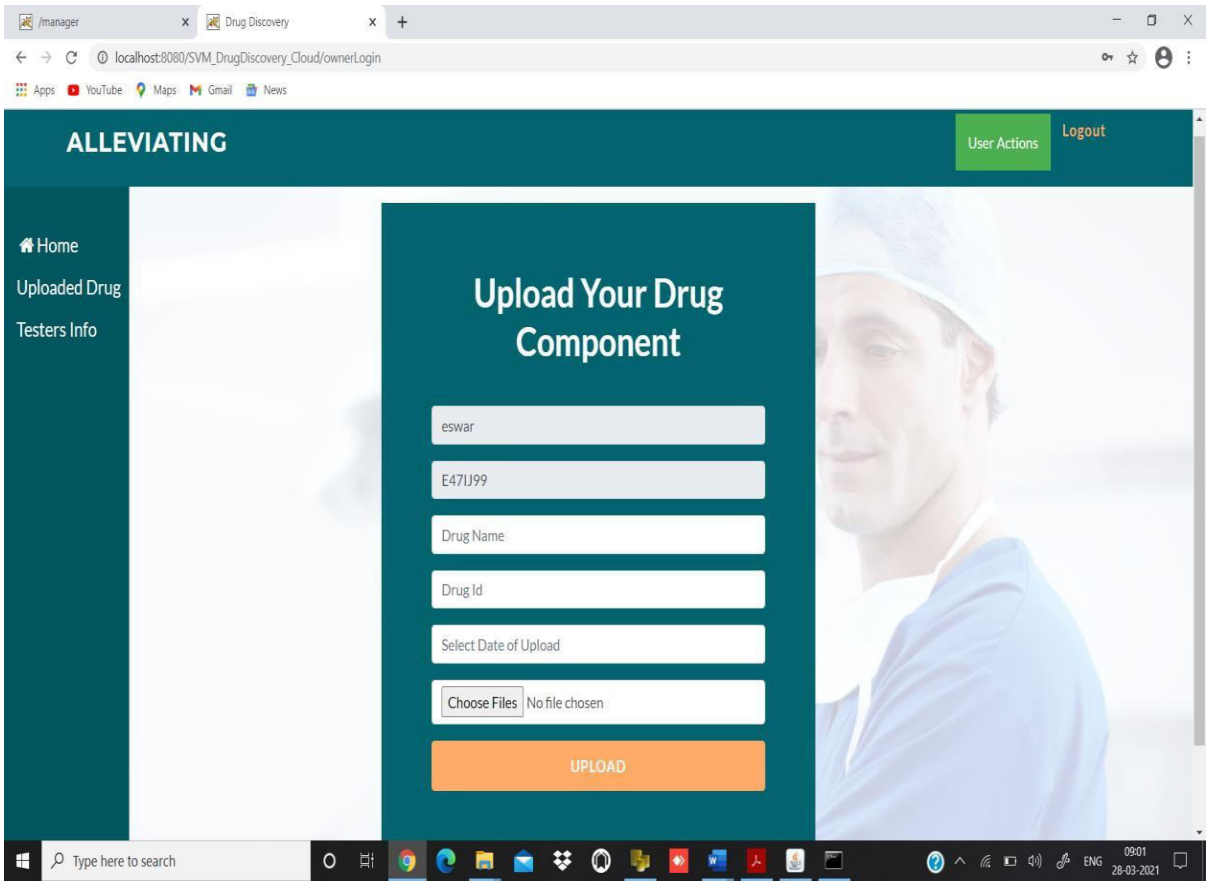


Fig.2 Uploading new Drug Components

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

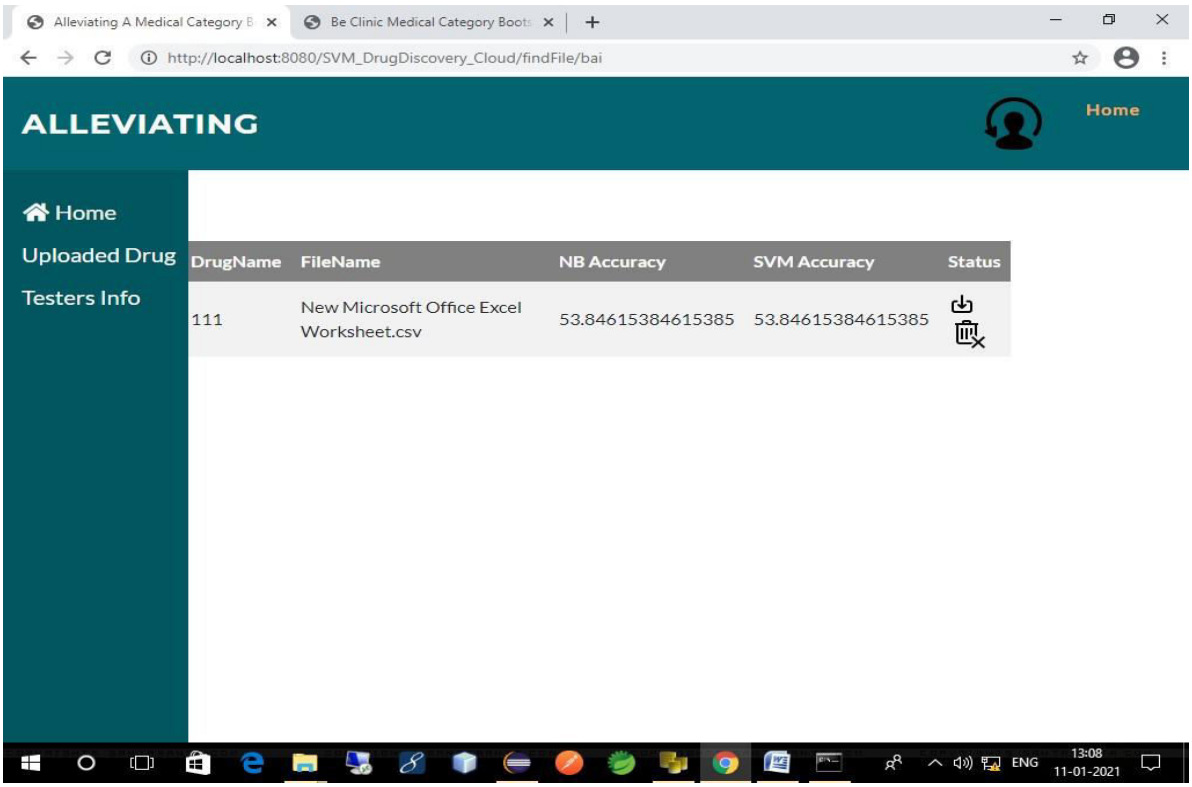


Fig.3 Checking Status

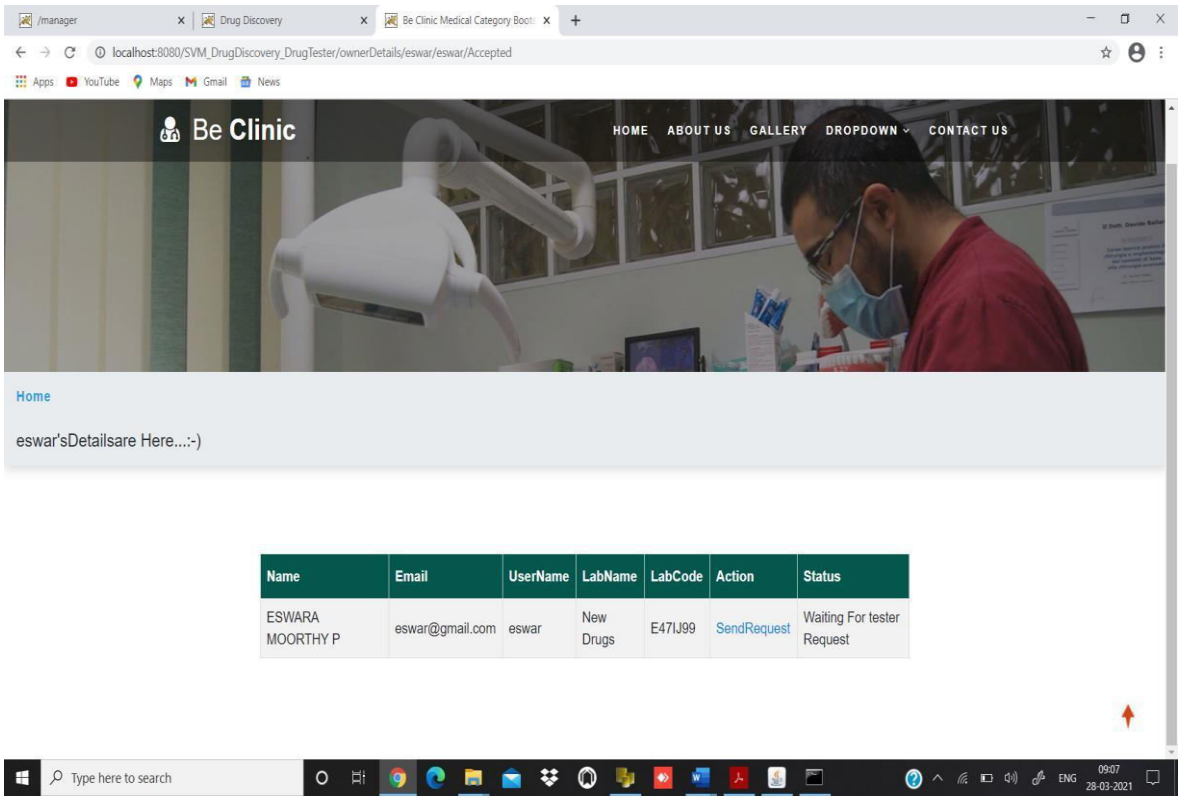


Fig.4 Testing the uploaded Drug Components

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

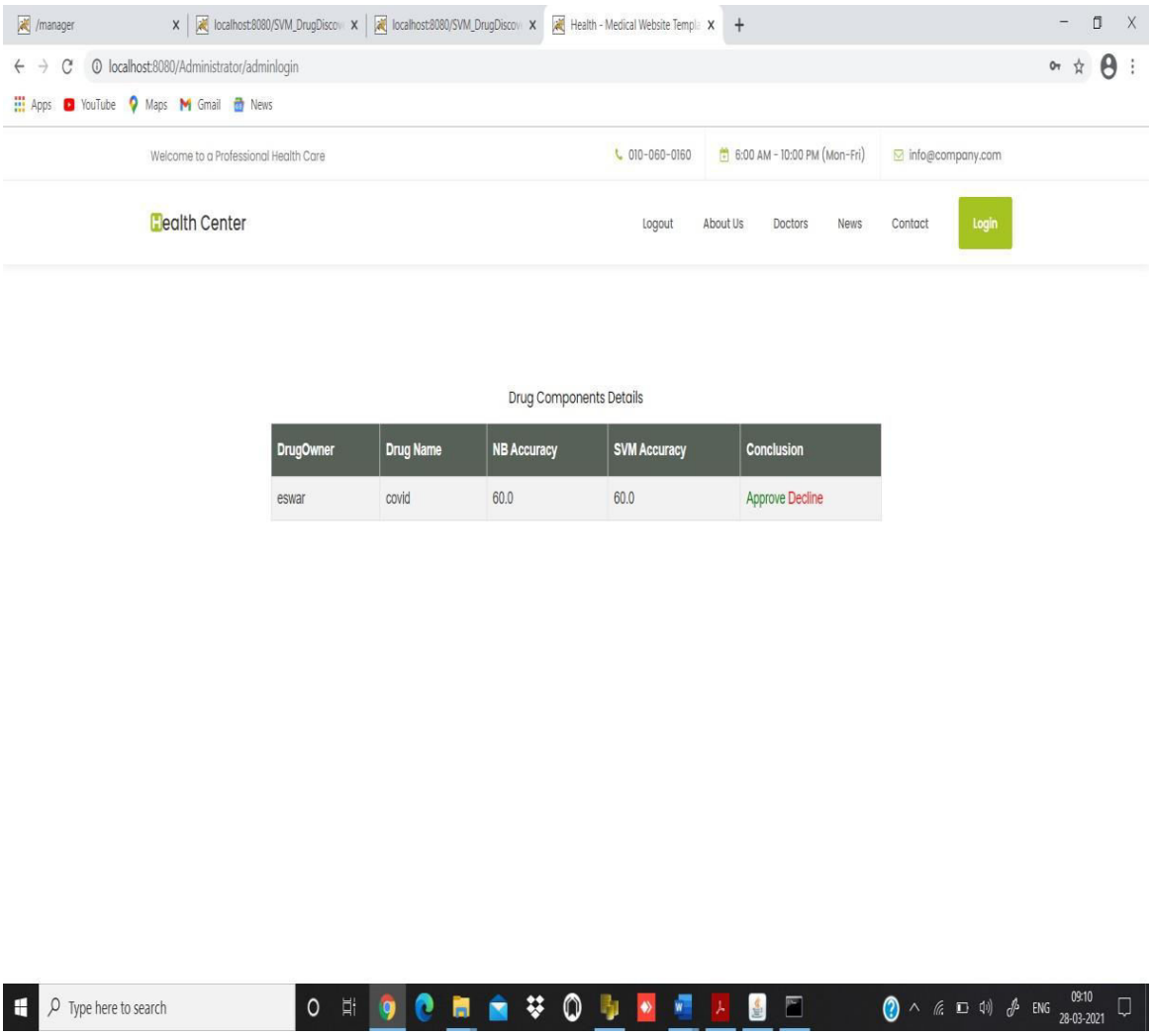


Fig.5 Admin Decision

V. CONCLUSION

In this Paper, We can check whether the new drug components are existing or not in secure way. The new drug components are checked with the existing datasets. The Screening test helps us to prevent unauthorized access to new Drug components. The time for getting the patent will be minimized and it will help the researchers to save time and they can get the process patent and continue with their research.

VI. FUTURE ENHANCEMENT

In the future work, we can improve our approach based on new benchmarks such that the accuracy and efficiency of the Screening test for the Drug Discovery Model. More number of Data can be handled effectively and those data can be used to check that are within medical standards. It will be more helpful for the Drug owners to speed up the process and more new drugs can be discovered at faster rate.

REFERENCES

[1] Kim-Kwang Raymond Choo;Ximeng Liu;Robert H. Deng;Yang Yang "Privacy-Preserving Outsourced Support Vector MachineDesign for Secure Drug Discovery",2020

[2]P. Hughes, S. Rees, S. B. Kalindjian, and K. L. Philpott, "Principles of early drug discovery," British journal of pharmacology, vol.162, no. 6, pp. 1239–1249, 2011.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903165 |

[4] I. Khanna, "Drug discovery in pharmaceutical industry: productivity challenges and trends," Drug discovery today, vol. 17, no. 19, pp.1088–1102, 2012.
[5] M. A. Lill and M. L. Danielson, "Computer-aided drug design platform using pymol," Journal of computer-aided molecular design,vol. 25, no. 1, pp. 13–19, 2011.
[6] "Research and markets, global drug discovery technologies market analysis & trends - industry forecast to 2025," http://www.researchandmarkets.com/research/n5klng/global drug.
[7] Y. Zhang and J. C. Rajapakse, Machine learning in bioinformatics. Wiley & Sons, 2009, vol. 4.
[8] J. B. Mitchell, "Machine learning methods in chemoinformatics," Wiley Interdisciplinary Reviews: Computational Molecular Science,
. 4, no. 5, pp. 468–481, 2014.
[9] T. Joachims, "Making large-scale svm learning practical," Technical Report,SFB 475: Komplexit¨atsreduktion in Multivariaten Datenstrukturen, Universit¨at Dortmund, Tech. Rep., 1998.
[10] R. Burbidge, M. Trotter, B. Buxton, and S. Holden, "Drug design by machine learning: support vector machines forpharmaceuticaldata analysis," Computers & chemistry, vol. 26, no. 1, pp. 5–14, 2001.
[11] R. Bost, R. A. Popa, S. Tu, and S. Goldwasser, "Machine learningclassification over encrypted data," in 22nd Annual Network and Distributed System Security Symposium, NDSS 2015, San Diego, California, USA, February 8-11, 2015, 2015.
[12] G. Cano, J. Garcia-Rodriguez, A. Garcia-Garcia, H. Perez-Sanchez, J. A. Benediktsson, A. Thapa, and A. Barr, "Automatic selection of descriptors using random forest: Application to drug discovery," Expert Systems with Applications, vol. 72, pp. 151–159,
[13] Liu, K.-K. R. Choo, R. H. Deng, R. Lu, and J. Weng, "Efficient and privacy-preserving outsourced computation of rational numbers," IEEE Journal of Biomedical and Health Informatics, vol. 20, pp. 655 – 668, 2016.
[14] X. Liu, R. Choo, R. Deng, R. Lu, and J. Weng, "Efficient and privacy-preserving outsourced calculation of rational numbers," IEEE Transactions on Dependable and Secure Computing, 2016.
[15] B. K. Samanthula, H. Chun, and W. Jiang, "An efficient and probabilisticbit decomposition," in Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security. ACM, 2013, pp. 541–546.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details