

## Chapter 8

# Deep Learning and Data Balancing Approaches in Mining Hospital Surveillance Data

**Adnan Firoze**

*North South University, Bangladesh*

**Tonmoay Deb**

*North South University, Bangladesh*

**Rashedur M. Rahman**

*North South University, Bangladesh*

### **ABSTRACT**

*A number of classifier models on hospital surveillance data to classify admitted patients according to their critical conditions with an emphasis to deep learning paradigms, namely convolutional neural network, were used in this research. Three class labels were used to distinguish the criticality of the admitted 25,261 patients. The authors have set forth two distinct approaches to address the unbalance nature of data. They used multilayer perceptron (MLP), convolutional neural network (CNN), and multinomial logistic regression classifications and finally compared the performance of our models with the models developed by Firoze, Hasan and Rahman (2013). After comparison, the authors show that one of the models, including convolutional neural network based on deep learning, surpasses most models in terms of classification performance in contingent with training times and epochs. The trade-off is computational power for which—to achieve optimal accuracy—multiple CUDA cores are necessary. The authors achieved stable improvement of classification for their model using CNN.*

## **1. INTRODUCTION**

Machine intervention in medicine and mining large scale medical surveillance data have caught significant attention in the recent years due to epidemics and the scarcity of physicians. We have pursued this research based on a dataset that stores patients' data from January 1, 1996 to December 31, 2007 (which is hospital surveillance data of 12 years) that was collected at International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B, 2008). Previously, a research work using this data repository was conducted using decision-tree induction algorithms by Rahman and Hasan (2011). We have introduced several newer approaches to deal the classification problem along with a novel way of balancing the dataset.

ICDDR,B established a diarrheal disease surveillance system in Dhaka, Bangladesh in 1979 and later extended it to their Matlab hospital at Comilla, Bangladesh in 2003. The surveillance system collects information on clinical, epidemiological and demographic characteristics of patients. A systematic 2% sub-sample of patients attending Clinical Research and Service Centre (CRSC) and all patients from the Health and Demographic Surveillance System (HDSS) area attending the Matlab hospital are enrolled into the surveillance program. The patients and/or their attendants supply information on socioeconomic and demographic characteristics, housing and environmental conditions, feeding practices, particularly among infants and young children, and on the use of drugs and fluid therapy at home to the interviewers. Moreover, nosocomial features e.g. clinical characteristics, anthropometric measurements, treatments received at the facility, and clinical outcomes of patients are also recorded. Extensive microbiological assessments of fecal samples (microscopy, culture, and ELISA) of patients are performed to identify diarrheal pathogens and to determine antimicrobial susceptibility of bacterial pathogens. It enables the center to detect the emergence of new pathogens and responds to early identification of outbreaks and their locations to suggest the Government of Bangladesh to take preventive measures.

Collected information is representative of the population and thus it serves as an important data repository for conducting epidemiological studies, validation of clinical studies, and it also helps develops new research ideas and study design.

### **1.1. Motivation**

Upon arrival at hospital, an initial diagnosis is carried out by the duty physician to find out the criticality of the patient's condition and upon completion, the duty doctor takes necessary action accordingly. This step becomes difficult yet more crucial in the event of an epidemic like that of the year when 1000 patient on an average got admitted to the hospital daily due to flood. The importance of this surfaced again in 2009 after the cyclone Aila hit the southern coast of Bangladesh. Similar picture has been drawn in USA during the recent Hurricane Hurvey. It becomes increasingly difficult to diagnose every patient satisfactorily due to scarcity of duty doctors. Thus, machine intervention to diagnose and measure the criticality of the newly arrived patient with the help of the historical data kept in the surveillance database was a necessity. The application asks few questions on physical condition and history of the patient and accordingly determines the critical condition of the patient as low, medium or high.

71 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/deep-learning-and-data-balancing-approaches-in-mining-hospital-surveillance-data/205125](http://www.igi-global.com/chapter/deep-learning-and-data-balancing-approaches-in-mining-hospital-surveillance-data/205125)

## Related Content

---

### A Software Modeling Perspective of UMLS Semantic Network

Rishi Kanth Saripalle (2018). *Handbook of Research on Emerging Perspectives on Healthcare Information Systems and Informatics* (pp. 305-322).

[www.irma-international.org/chapter/a-software-modeling-perspective-of-umls-semantic-network/205131](http://www.irma-international.org/chapter/a-software-modeling-perspective-of-umls-semantic-network/205131)

### Managed Healthcare: Doctor Life Satisfaction and Its Impact on the Process of Communicating With the Patient

Anna Rosiek and Aleksandra Rosiek-Kryszewska (2018). *Healthcare Administration for Patient Safety and Engagement* (pp. 244-261).

[www.irma-international.org/chapter/managed-healthcare/197564](http://www.irma-international.org/chapter/managed-healthcare/197564)

### Economic Profit vs. Social Benefit: An Approach to the Pharmaceutical Industry

M. Mercedes Galán-Ladero and M. Ángeles Galán-Ladero (2021). *Management and Marketing for Improved Competitiveness and Performance in the Healthcare Sector* (pp. 24-44).

[www.irma-international.org/chapter/economic-profit-vs-social-benefit/285800](http://www.irma-international.org/chapter/economic-profit-vs-social-benefit/285800)

### Drawing Fragmented Narratives: Alzheimer's Through the Looking Glass

Chiara Battisti (2023). *Global Perspectives on Probing Narratives in Healthcare* (pp. 170-190).

[www.irma-international.org/chapter/drawing-fragmented-narratives/324293](http://www.irma-international.org/chapter/drawing-fragmented-narratives/324293)

### The Potential of System Dynamics to Model Patient-Aided Healthcare

(2020). *Managing Patients' Organizations to Improve Healthcare: Emerging Research and Opportunities* (pp. 93-115).

[www.irma-international.org/chapter/the-potential-of-system-dynamics-to-model-patient-aided-healthcare/246996](http://www.irma-international.org/chapter/the-potential-of-system-dynamics-to-model-patient-aided-healthcare/246996)