

## OFFICIAL ABSTRACT and CERTIFICATION

### Improving Hepatocellular Carcinoma Survival Prediction with Artificial intelligence Strategies

Lingfei Zhao

Syosset High School, Syosset, NY, US

Hepatocellular Carcinoma (HCC), which constitutes over 90 percent of liver cancers, is the sixth most frequently diagnosed cancer and the third-leading cause of cancer related deaths in the world. Its incidence rate has tripled since 1980 and is estimated to further increase 61.9% within two decades. Thus, the increasing prevalence of HCC emphasizes the need for more accurate methods of survival prediction in order to increase survival rate.

Two proposed methods of survival prediction for Hepatocellular Carcinoma patients were compared, both utilizing the missForest algorithm to impute missing values in the dataset. The first method incorporated random forests to classify the data into two categories: Positive (the patient will not survive past one year) and Negative (the patient will survive past one year), while the second method used a combination of principal component analysis and support vector machine (PCA-SVM) to make these predictions. Although the Random Forest model had a higher mean accuracy of 0.723 compared with the PCA-SVM model's mean accuracy of 0.712, the latter had a significantly higher mean True Positive Rate (0.630 compared with the 0.547 of the Random Forest model,  $p < 0.05$ ). In other words, the PCA-SVM model was superior in correctly predicting that a patient was in critical condition and would not survive past one year. This model allows for more effective treatment to target the patient's individual needs, thus, the PCA-SVM was recommended as a more reliable method for prognosis evaluation.

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check ALL that apply):
 

☐ human participants

☐ potentially hazardous biological agents

☐ vertebrate animals

☐ microorganisms

☐ rDNA

☐ tissue
2. I/we worked or used equipment in a regulated research institution or industrial setting: ☐ Yes ☒ No
3. This project is a continuation of previous research. ☐ Yes ☒ No
4. My display board includes non-published photographs/visual depictions of humans (other than myself): ☐ Yes ☒ No
5. This abstract describes only procedures performed by me/us, reflects my/our own independent research, and represents one year's work only: ☒ Yes ☐ No
6. I/we hereby certify that the abstract and responses to the above statements are correct and properly reflect my/our own work. ☒ Yes ☐ No

*This stamp or embossed seal attests that this project is in compliance with all federal and state laws and regulations and that all appropriate reviews and approvals have been obtained including the final clearance by the Scientific Review Committee.*

Category

Pick one only — mark an "X" in box at right

Animal Sciences

Behavioral & Social Sciences

Biochemistry

Biomedical & Health Sciences

Biomedical Engineering

Cellular & Molecular Biology

Chemistry

Computational Biology & Bioinformatics

Earth & Environmental Sciences

Embedded Systems

Energy: Sustainable Materials and Design

Engineering Mechanics

Environmental Engineering

Materials Science

Mathematics

Microbiology

Physics & Astronomy

Plant Sciences

Robotics & Intelligent Machines

Systems Software

Translational Medical Sciences

