World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

A Machine Learning-Based Model to Screen Antituberculosis Compound Targeted against LprG Lipoprotein of Mycobacterium tuberculosis

Authors : Syed Asif Hassan, Syed Atif Hassan

Abstract : Multidrug-resistant Tuberculosis (MDR-TB) is an infection caused by the resistant strains of Mycobacterium tuberculosis that do not respond either to isoniazid or rifampicin, which are the most important anti-TB drugs. The increase in the occurrence of a drug-resistance strain of MTB calls for an intensive search of novel target-based therapeutics. In this context LprG (Rv1411c) a lipoprotein from MTB plays a pivotal role in the immune evasion of Mtb leading to survival and propagation of the bacterium within the host cell. Therefore, a machine learning method will be developed for generating a computational model that could predict for a potential anti LprG activity of the novel antituberculosis compound. The present study will utilize dataset from PubChem database maintained by National Center for Biotechnology Information (NCBI). The dataset involves compounds screened against MTB were categorized as active and inactive based upon PubChem activity score. PowerMV, a molecular descriptor generator, and visualization tool will be used to generate the 2D molecular descriptors for the actives and inactive compounds present in the dataset. The 2D molecular descriptors generated from PowerMV will be used as features. We feed these features into three different classifiers, namely, random forest, a deep neural network, and a recurring neural network, to build separate predictive models and choosing the best performing model based on the accuracy of predicting novel antituberculosis compound with an anti LprG activity. Additionally, the efficacy of predicted active compounds will be screened using SMARTS filter to choose molecule with drug-like features.

Keywords : antituberculosis drug, classifier, machine learning, molecular descriptors, prediction
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020