

Título de la Tesis: “Técnicas de aprendizaje automático y computación científica aplicadas a la predicción de parámetros ADME-Tox”

Doctorado en Ciencias de la Computación

Autor: Soto, Axel Juan

Directores: Dr. Ignacio Ponzoni - Dr. Gustavo Vázquez

Resumen

Hace 15 años atrás, el desarrollo de nuevos productos farmacéuticos consistía en un proceso de prueba y error basado mayormente en la búsqueda de un farmacóforo o principio activo. Sin embargo, muchos compuestos eran finalmente descartados en la última etapa del proceso de su desarrollo debido al comportamiento ADME-Tox. Por estos motivos, el interés en la industria y el ámbito académico en las disciplinas de quimioinformática y, en particular, en las técnicas de tipo QSAR ha crecido considerablemente en los últimos años.

El objetivo de esta tesis se centró en desarrollar metodologías para la mejora de los modelos existentes de QSAR mediante el uso de técnicas numéricas y de aprendizaje automático. En este sentido, se obtuvo un potente método para lo que se considera uno de los problemas más importantes del modelado QSAR: la selección de subconjuntos de descriptores relevantes. Para esta tarea se desarrollaron distintos enfoques utilizando computación evolutiva.

Asimismo, otro aspecto central considerado fue el desarrollo de una técnica de identificación de dominio de aplicación para un método de predicción, el cual permite determinar los alcances en las predicciones de un modelo. Para esta técnica se consideró la aplicación de medidas de distancia entre compuestos químicos, usando aprendizaje no supervisado. Finalmente, se desarrolló un método generalizado que permite la proyección de los datos en un espacio de menor dimensión, en donde las distancias entre los datos proyectados guardan relación con las distancias de la propiedad a modelar. Este nuevo espacio permite mejorar la visualización, reducir el conjunto de descriptores en forma embebida y mejorar la precisión de los modelos de predicción.

Título de la Tesis: “Técnicas de aprendizaje automático y computación científica aplicadas a la predicción de parámetros ADME-Tox”

Doctorado en Ciencias de la Computación

Autor: Soto, Axel Juan

Directores: Dr. Ignacio Ponzoni - Dr. Gustavo Vázquez

Abstract

15 years ago, development of new drugs was based on a trial-and-error process that was mostly devoted to the searching of pharmacophore fragments or drug potency. Nevertheless, many compounds were discarded in the latest stages of their development due to poor ADME-Tox behavior. Therefore, chemoinformatics interest given by the scientific and industrial community has grown considerably in the last years.

The thesis' objective is focused on the development of statistical and machine learning techniques aimed at improving current QSAR-based methods' limitations. In this way, a robust method for tackling the problem of selecting relevant descriptor subsets was developed. Selection of descriptors is one of the most important challenges in QSAR. This task was carried out by means of evolutionary computing techniques.

Moreover, another crucial issue was the development of a method for identifying the applicability domain of a given prediction method, in order to estimate the scope of the accuracy of a model. Similarity metrics and unsupervised learning were applied for this task. Finally, a general approach for data projection onto a low-dimensional space was proposed, where distances among projected data are in maximum correlation with distances on the target space. This new subspace projection allows a better visualization capacity, an embedded descriptor selection and an improvement of the prediction capacity.