

Implementación de redes neuronales convolucionales para el estudio de interacciones proteína-ligando

Adrián Antonio Rodríguez Pié

El descubrimiento de nuevos fármacos es un proceso muy tardado y costoso. El desarrollo e incluso el reposicionamiento de compuestos ya conocidos es una tarea sumamente complicada. El escenario es aún más complicado si se toman en cuenta los miles o millones de moléculas capaces de ser sintetizadas en cada etapa del desarrollo. Para superar estas dificultades, el uso de alternativas computacionales de bajo costo es altamente recomendado, y se ha adoptado como la forma estándar de ayuda para el desarrollo de nuevos fármacos.

Una de las metodologías computacionales más usadas para investigar estas interacciones es el acoplamiento molecular. La selección de los ligandos más potentes utilizando Cribado Virtual basado en Acoplamiento (DBVS) es realizado a través de realizar la inserción de cada compuesto de la biblioteca de compuestos en una región particular de un receptor objetivo. En la primera etapa del proceso, una búsqueda heurística es llevada a cabo en la que miles de posibles inserciones son consideradas. En la segunda etapa, la calidad de la inserción es evaluada a través de una función evaluadora. Esta última fase se ha convertido en todo un reto para los científicos computacionales, por la dificultad de decidir de forma determinística si un acoplamiento es bueno o no.

Los sistemas basados en aprendizaje de máquina (ML) han sido usados con éxito para mejorar la salida del DBVS para tanto incrementar el desempeño de las funciones evaluadoras, como para construir clasificadores de afinidad de enlace. Una de las principales ventajas de utilizar ML es la capacidad de explicar la dependencia no lineal de las interacciones moleculares entre ligando y receptor.

En este trabajo se propone un acercamiento con una red neuronal convolucional para mejorar el DBVS. El método utiliza los resultados de una simulación de acoplamiento como entrada, en donde automáticamente aprende a extraer características relevantes a partir de datos básicos como tipos de átomo, distancias entre ellos, y su contexto en la rama estructural. La red aprende características abstractas que son útiles para discriminar entre ligandos activos y señuelos en una proteína.

Bibliografía

- [1] ARCINIEGA M, & LANGE OF. Improvement of Virtual Screening Results by Docking Data Feature Analysis. *J. Chem. Inf. Model.* **2014**; 54:1401-11.
- [2] LOUNNAS V, RITSCHER T, KELDER J, MCGUIRE R, BYWATER RP, & FOLOPPE N. Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. *Comput. Struct. Biotechnol. J.* **2013**; 5:e201302011.
- [3] PEREIRA, J. C., CAFFARENA, E. R., & DOS SANTOS, C. N. Boosting docking-based virtual screening with deep learning. *Journal of chemical information and modeling* **2016**; 56:2495-2506.
- [4] BISHOP, C. (2007). *Pattern Recognition and Machine Learning (Information Science and Statistics)*, 1a edición. 2006. corr. 2a edición. Springer, New York.
- [5] JOHN LU, Z. Q. (2010). *The elements of statistical learning: data mining, inference, and prediction*. Journal of the Royal Statistical Society: Series A (Statistics in Society), 173:693-694.

Tesista

Asesor