

Identificación del proyecto

Nombre del proyecto

Procesado de señal en metabolómica cuantitativa dirigida y descubrimiento de biomarcadores volátiles basado en aprendizaje automático: aplicación a cribado de cáncer

Expediente numero

PID2021-126543OB-C22



Descripción del proyecto

La metabolómica es la última de las ciencias ómicas que han aparecido para intentar implementar la medicina personalizada. Su objetivo es caracterizar completamente el conjunto de moléculas pequeñas presentes en un biofluido (el metaboloma). Una de las principales ventajas de la metabolómica es que mientras otras ciencias "ómicas" pueden decirte lo que puede suceder, conocer el metaboloma de un organismo vivo te dice lo que realmente está sucediendo.

La hipótesis subyacente de la metabolómica del cáncer es que la composición química de la mayoría de los biofluidos accesibles en humanos se altera de formas desconocidas. Para caracterizar esos cambios, las mediciones en biofluidos humanos con instrumentación química de última generación proporcionan una gran cantidad de datos sin procesar que son difíciles de interpretar clínicamente.

La metabolómica computacional no orientada tiene como objetivo extraer la mayor cantidad de información posible de una gran cantidad de datos sin procesar y traducir esa información en biomarcadores clínicos para el diagnóstico de enfermedades. Esto depende en gran medida del uso de procesamiento de señales y diferentes tipos de algoritmos de aprendizaje automático. Existe una necesidad no cubierta consistente en el desarrollo de flujos de trabajo de procesamiento de datos para analizar datos brutos cuando se utiliza cromatografía de gases y espectrometría de masas.

Pero incluso si se puede hacer una distinción y se encuentran biomarcadores, en la mayoría de los casos se desconoce la identidad química del biomarcador. Ésta es una de las principales dificultades para trasladar estos resultados al ámbito clínico. En la mayoría de los casos, los médicos no solo están interesados en modelos predictivos, sino también en la interpretación biológica de los biomarcadores encontrados.

Por esta razón, nuestro objetivo en este proyecto es utilizar información de la base de datos del metaboloma humano para desarrollar una biblioteca extensa de compuestos clave en la orina humana que están asociados con importantes procesos metabólicos. Aquí, el mayor desafío es identificar y cuantificar compuestos específicos (un enfoque dirigido) dentro de la miríada de compuestos presentes en el biofluido seleccionado: la orina.

De esta manera, esperamos poder cuantificar un conjunto dado de metabolitos clave, fijos y específicos con identidades químicas previamente conocidas.

Incluso con un enfoque dirigido, el uso de métodos de aprendizaje automático para el análisis de datos y la selección de características sigue siendo una gran necesidad si el perfil de metabolitos identificado tiene una alta cardinalidad. Además, un subconjunto de enfoques de aprendizaje automático, las técnicas de aprendizaje profundo están ganando mucho terreno para el pre-procesamiento de datos a bajo nivel, como la detección e identificación de picos.

En este proyecto queremos combinar esta clasificación de características con un algoritmo de búsqueda adecuado para la selección final del mejor subconjunto para los modelos de predicción.

Un aspecto destacado de este proyecto es el uso de varias técnicas de análisis (con diferentes instrumentos de última generación) en alícuotas de las mismas muestras. Esperamos encontrar información complementaria sobre la composición de la muestra y, por lo tanto, vale la pena investigar las técnicas de integración de datos multiómica.

Financiación

Entidad financiadora

MCIN/ AEI /10.13039/501100011033/ y por FEDER Una manera de hacer Europa

Importe

181.500,00 €