

INFORME DEL PROYECTO REF.SV-24-GIJÓN-1-16

(AI4PSA) Modelado de la evolución del biomarcador PSA mediante técnicas IA para la predicción de la recidiva de cáncer de próstata

Fechas inicial y final del proyecto:

01 / 01 / 2024 al 31 / 12 / 2024

Investigador/a Principal:

Víctor Manuel González Suárez

Otros investigadores:

Héctor Gómez

Universidad de Purdue

Personal contratado:

José Ramón Suárez Bajo

Fechas inicial y final de contratación:

01 / 09 / 2024 al 31 / 12 / 2024

Empresas o instituciones colaboradoras:

Universidad de Purdue

Instituto de Investigación Sanitaria de Santiago de Compostela

Redes sociales de investigadores y empresas:

<https://www.researchgate.net/profile/Victor-Gonzalez-97>

<https://www.linkedin.com/in/v%C3%ADctor-m-gonz%C3%A1lez-224a832a/>

https://engineering.purdue.edu/ME/People/ptProfile?resource_id=157698

<https://www.researchgate.net/profile/Hector-Gomez-5>

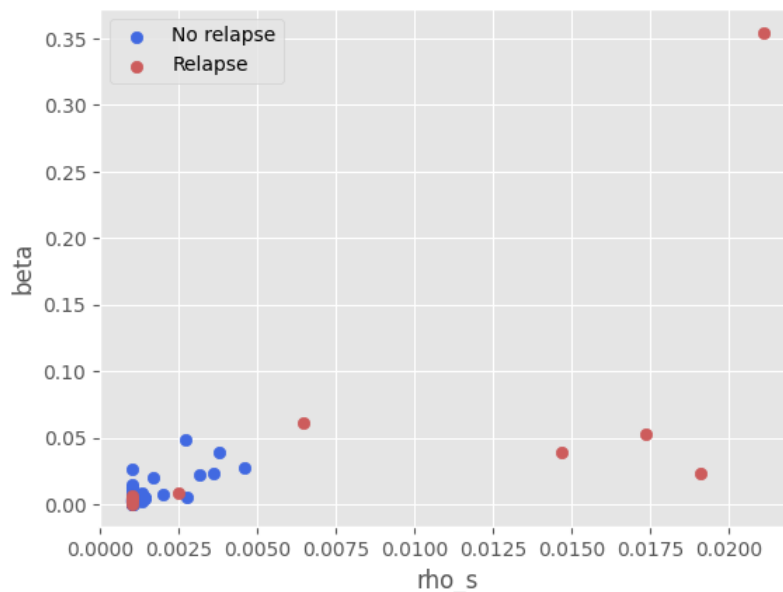
<https://www.linkedin.com/in/hector-gomez-7b14971b8/>

<https://www.idisantiago.es/gl/grupo-de-investigacion/e043-fisica-medica-y-biomatematicas>

Resumen Gráfico

Añade una imagen, gráfico o diagrama de Gantt que resuma tu proyecto. **No se permite** el uso de varias imágenes superpuestas entre sí ya que puede ocasionar un problema serio debido a que pueden moverse al pasarlo de un archivo word / odt / rtf de un dispositivo a otro.

Memoria descriptiva del proyecto



1. Resumen ejecutivo

Antecedentes

Según el Instituto Nacional del Cáncer de USA [1], el Prostate Specific Antigen (antígeno prostático específico) o PSA es una proteína que producen tanto las células normales como las células malignas (cancerosas) de la próstata. La prueba del PSA se usa para medir la concentración del PSA en la sangre.

El PSA se considera un buen biomarcador del cáncer de próstata ya que hay una correlación directa entre la concentración elevada de PSA en sangre y dicha enfermedad. En 1986, la Administración de Alimentos y Medicamentos de USA (FDA) aprobó por primera vez la prueba de PSA para vigilar el avance del cáncer de próstata en los hombres que ya tenían un diagnóstico de la enfermedad. En 1994, la FDA aprobó el uso de la prueba del PSA para acompañar el Examen Digital del Recto (EDR) y ayudar a detectar el cáncer de próstata en los hombres de 50 o más años. A partir más o menos de 2008, muchos médicos y organizaciones profesionales recomendaban el uso de esta prueba en la detección sistemática anual del cáncer de próstata a partir de los 50 años.

Se han realizado varios estudios a nivel de Estados Unidos, de la UE y de Reino Unido sobre la incidencia del cáncer de próstata y el teste de PSA como mecanismo diagnóstico. Se ha demostrado que su uso tiene ciertas limitaciones en el diagnóstico inicial de la enfermedad, pero al mismo tiempo se ha demostrado que, la evolución de la PSA, en aquellos pacientes que han tenido la enfermedad y han sido sometidos a tratamiento de radioterapia, es un buen biomarcador de la probabilidad de recidiva de esta enfermedad.

Si la concentración del PSA de un hombre comienza a aumentar después del tratamiento de cáncer de próstata, podría ser el primer signo de recidiva. Esto se llama "recidiva bioquímica" y suele aparecer meses o años antes de que la recidiva cause síntomas.

Esta capacidad de anticipación de este biomarcador es lo que hace que sea tan interesante el controlar su evolución a lo largo del tiempo.

Sin embargo, la evolución de la PSA no es igual en todos los pacientes, sino que presenta diferencias sustanciales de unos a otros. De ahí que sea necesario contar con herramientas que permitan tener modelos de la evolución de PSA personalizados para cada paciente, que ayuden a los especialistas médicos a diagnosticar de forma más precisa la presencia o no de recidiva.

En este sentido, la Inteligencia Artificial (IA) puede ser una herramienta muy útil para desarrollar ese tipo de modelos de evolución de la PSA a medida de cada paciente.

Objetivo general

El objetivo general de este proyecto consistió en tratar de acortar el tiempo de detección de una recidiva química de pacientes de cáncer de próstata sometidos a un tratamiento de radioterapia. Para ello se continuó el trabajo iniciado y no culminado en el proyecto del año 2023, consistente en desarrollar y validar algoritmos de IA que permitan obtener modelos personalizados de evolución de la PSA en este tipo de pacientes.

Metodología

La experimentación planteada constó de varias etapas:

1. Aplicar algoritmos de detección de outliers y el análisis del grado de acierto.
2. Estudio de modelado y clasificación "on-line" basado en estudios previos, integración de árboles de decisión.
3. Estudiar la combinación de métodos de detección de anomalías, agrupamiento y clasificación.

Resultados

Los resultados obtenidos en este proyecto no son los esperados. Primeramente, el análisis de outliers no ha generado un reconocimiento de adecuado de casos de recaída, con lo que se ha desechado continuar con el uso de estos datos, al menos, mientras no se disponga de más datos. La realidad indica que con los algoritmos de detección de outliers no mejoran los resultados presentados por el modelo mecanicista.

Por otra parte, el uso de un conjunto de datos de entrenamiento, asumiendo que los datos de validación se utilizan para entrenamiento o enriquecimiento posterior de los modelos (la aproximación online propuesta) tampoco ha generado unos datos que permitan aumentar el nivel significativo de clasificación de los modelos mecanicistas.

Consecuentemente, la última etapa de la metodología no se desarrolló por carecer de sentido, dado que claramente los resultados iban a ser insatisfactorios.

Trabajos futuros

No obstante lo comentado anteriormente, la investigación no ha terminado con este proyecto. Los resultados obtenidos inducen a pensar en modelos neuronales profundos serán los apropiados para la mejora del nivel de acierto del modelo mecanicista. Todo ello sería posible fusionando conjuntos de datos provenientes de diferentes data sets, en particular, los dos conjuntos de datos disponibles en este contexto. Cabe añadir, que se podría optar por modelos supervisados y no supervisados, como por ejemplo, autoencoders para transformar el dominio, y redes profundas para la clasificación. Esta propuesta sería una posible línea de continuación en futuras convocatorias.

2. Objetivos iniciales del proyecto y grado de consecución

Tipo	Objetivo	Comentarios	Grado de consecución
Principal	Acortar el tiempo de detección de una recidiva de pacientes de cáncer de próstata que hayan sido sometidos a un tratamiento de radioterapia. Para ello se desarrollarán y validarán algoritmos de IA que permitan obtener modelos personalizados de evolución de la PSA en este tipo de pacientes	Se han explorado diferentes métodos pero no se han logrado resultados positivos. Se deberían probar otros métodos diferentes.	80%
Secundario/ Específico	Explorar distintas posibilidades de combinar los resultados de la búsqueda de "outliers" en cada una de las variables para alcanzar un resultado final más robusto	Con métodos clásicos de detección de outliers	100%
Secundario/ Específico	Analizar los resultados que nos da cada uno de los clasificadores del sistema desarrollado hasta ahora por separado.	Combinando estudios previos y detección de anomalías	100%
Secundario/ Específico	Explorar la eficacia de los clasificadores del tipo árboles de decisión.	Analizado en el contexto anterior	100%
Secundario/ Específico	Explorar los resultados que da nuestro sistema utilizando como guía otras medidas de la calidad del agrupamiento. Con las pruebas que tenemos, el agrupamiento guiado por el "Silhouette Score" parece un buen candidato.	Diseño de método que selecciona el mejor agrupamiento utilizando diferentes métricas de calidad de agrupamiento	100%

Secundario/ Específico	Aplicar esta detección de valores atípicos para mejorar el sistema de clasificadores.	Los resultados no fueron los esperados	100%
Secundario/ Específico	Estudiar cómo ensamblar los datos que nos proporcionan ambos métodos.	Debido a los resultados, se optó por no desarrollar esta etapa	0 %

3. Tareas realizadas:

ID	Descripción	MESES					
		1	2	3	4	5	6
Fase 1	Mejora de la robustez de los algoritmos de búsqueda de outliers.						
Tarea 1.1	Evaluación de la combinación de resultados de los algoritmos de búsqueda de outliers desarrollados.						
Fase 2	Clasificadores						
Tarea 2.1	Análisis en detalle del funcionamiento de los clasificadores desarrollados.						
Tarea 2.2	Implementación de nuevos clasificadores.						
Tarea 2.3	Combinación de los resultados de los clasificadores.						
Fase 3	Medidas de calidad de agrupamiento.						
Tarea 3.1	Análisis y clasificación de las métricas de calidad de agrupamientos.						
Tarea 3.2	Implementación de Silhouette Score.						
Fase 4	Modelo de predicción de recidiva química.						
Tarea 4.1	Modelado del sistema de predicción.						
Tarea 4.2	Integración de los módulos de búsqueda de outliers, clasificadores y medidores de calidad del agrupamiento.						
Fase 5	Desarrollo de publicaciones científicas						
Tarea 5.1	Publicación en revistas y congresos de los resultados obtenidos						

4. Resultados obtenidos:

Los resultados obtenidos en la mejora de la robustez de los algoritmos de búsqueda de valores atípicos (outliers) se pueden considerar insatisfactorios. Básicamente, no se ha conseguido mejorar los intervalos de confianza del modelo mecanicista. En principio, aplicando técnicas clásicas de detección de outliers, así como métodos de detección de anomalías, no han conseguido incrementar el número de aciertos sobre las instancias en los conjuntos de validación.

De igual manera, las aproximaciones secuenciales, pensadas para solucionar el problema de escasez de datos válidos, se han mostrado poco robustas, altamente dependientes de los datos disponibles en cada momento.

Las principales consecuencias de estos resultados son que o bien el diseño experimental no consideró algún aspecto relevante -con lo cuál, habría que proceder a analizar el diseño realizado y a su mejora-, o los modelos empleados no son capaces de discriminar adecuadamente las instancias positivas de entre las negativas.

A partir de estas conclusiones, es posible pensar en el uso de modelos neuronales profundos para i) convertir el espacio de dominio en un espacio latente más representativo, ii) aplicar modelos profundos -conjuntamente con la fusión de los data sets abiertos disponibles para aumentar el número de instancias- para discriminar mejor entre las instancias positivas y negativas, aumentando la potencialidad del modelo mecanicista.

5. Trabajos o necesidades futuras

Como ya se ha indicado anteriormente, las necesidades futuras pasan por utilizar de modelos neuronales profundos para i) convertir el espacio de dominio en un espacio latente más representativo, ii) aplicar modelos profundos -conjuntamente con la fusión de los data sets abiertos disponibles para aumentar el número de instancias- para discriminar mejor entre las instancias positivas y negativas, aumentando la potencialidad del modelo mecanicista.

6. Divulgación de los resultados

Debido a los malos resultados obtenidos por las propuestas, la difusión de resultados no ha sido la esperada. En concreto, no se ha publicado en ningún congreso ni revista los mismos. Sin embargo, sí se ha difundido la investigación realizada a través de una charla en los desayunos tecnológicos, en concreto, en la jornada del 8 de noviembre de 2024 titulada *Aplicación de IA al análisis de evolución del antígeno PSA para modelar la evolución del cáncer de próstata* organizada por el Instituto Universitario de Tecnología Industrial de Asturias (IUTA) en combinación con Gijón Impulsa.

Memoria económica:

1. Gastos:

Añade los gastos finales que has tenido a lo largo del proyecto.

Concepto	Gasto
----------	-------

Personal (IUTA)	12.000€
Fungibles	300€
Amortización	1.000€
Otros (Desplazamientos, Inscripciones a Congresos, etc.)	6.000€
TOTAL GASTOS	19.300€

2. Ingresos:

Añade los ingresos finales que has tenido a lo largo del proyecto. Indica las empresas y las referencias de los proyectos/contratos.

Entidad/Empresa financiadora Ref. Proyecto/Contrato	Concepto	Ingreso
IUTA	Ayuda IUTA: Contratación de personal	2.600 €
Otras entidades / empresas financiadoras		0 €
Financiación propia		16.700€
TOTAL INGRESOS		19.300 €

Bibliografía:

1. Instituto Nacional del Cáncer de Estados Unidos de América. Instituto National Institutes of Health (NIH).
<https://www.cancer.gov/espanol/tipos/prostata/hoja-informativa-psa#se-recomienda-la-prueba-del-psa-para-detectar-el-cancer-de-prostata>
2. Patrick A. Kupelian a, Arul Mahadevan b, Chandana A. Reddy b, Alwyn M. Reuther b, <nd Eric A. Klein c; "Use of different definitions of biochemical failure after external beam radiotherapy changes conclusions about relative treatment efficacy for localized prostate cancer". Urology. Volume 68, Issue 3, September 2006, Pages 593-598. Elsevier.
3. Eli Rosenbaum, Alan Partin, and Mario A. Eisenberger. "Biochemical Relapse After Primary Treatment for Prostate Cancer: Studies on Natural History and Therapeutic Considerations". Journal of the National Comprehensive Cancer Network. Volume 2: Issue 3. May 2004.
4. Jie Chen, C. Daniel Mullins, Priscilla Novak, and Stephen B. Thomas. "Personalized Strategies to Activate and Empower Patients in Health Care and Reduce Health Disparities". Health Education Behaviour. 2016 Feb; 43(1): 25–34. National Library of Medicine.
5. Krzyszczyk P, Acevedo A, Davidoff EJ, Timmins LM, Marrero-Berrios I, Patel M, White C, Lowe C, Sherba JJ, Hartmanshenn C, O'Neill KM, Balter ML, Fritz ZR, Androulakis IP, Schloss RS, Yarmush ML. The growing role of precision and personalized medicine for cancer treatment. Technology (Singap World Sci). 2018 Sep-Dec;6(3- 4):79-100. doi: 10.1142/S2339547818300020. Epub 2019 Jan 11. PMID: 30713991; PMCID: PMC6352312.