



1. DATOS BÁSICOS DEL TFG:

Título: Separación de señal de muones y piones en el experimento SBND mediante técnicas de aprendizaje supervisado

Descripción general (resumen y metodología):

La identificación de patrones y elaboración de predicciones está presente en muchas aplicaciones de la Física. Distintos algoritmos, basados en modelos probabilísticos, son la base del aprendizaje supervisado. Estos modelos permiten un aprendizaje previo basado en un etiquetado de los datos, que permite tomar decisiones o hacer predicciones en función de un conjunto de variables explicativas. La identificación del carácter cuantitativo o cualitativo de las variables de entrada, la distribución de probabilidad conjunta de las variables explicativas, etc., es el punto de partida para la elección de una metodología de clasificación adecuada [1]. Uno de los contextos en los que dichas técnicas se han usado con mayor éxito hasta la fecha corresponde al análisis de los sucesos registrados por detectores de Física de partículas. En este trabajo se propone aplicar técnicas de aprendizaje supervisado y análisis multivariable en el contexto del experimento de neutrinos SBND [5] [6]. Este experimento, cuyo objetivo primordial es confirmar o refutar la existencia de un cuarto neutrino estéril empleando para ello cámaras de deriva de argón líquido, cuenta con un programa científico mucho más amplio. En particular, podrá realizar una medición muy precisa de la sección eficaz en colisiones de neutrinos con blanco fijo, tanto en los canales inclusivos como en topologías muy específicas, gracias a la ingente estadística de sucesos de que dispondrá. Tanto para dichas mediciones, como para el análisis principal de búsqueda del neutrino estéril [7], es crucial disponer de una clasificación eficiente de las partículas en el estado final. En este trabajo se pretende que el alumno, a partir de una discusión general sobre la problemática de la clasificación estadística, profundice en el análisis del modelo más adecuado según el carácter de las variables de entrada, dando una visión completa y actualizada de sus aspectos fundamentales, su implementación, su validación y su aplicación a la separación de partículas en el estado final en el experimento SBND. Tradicionalmente, el proceso de clasificación de la señal de las partículas cargadas se basa por lo general en su distinta deposición de energía a medida que se desplazan por el argón. Sin embargo, la similitud entre la masa del muon (105.7 MeV) y del pion cargado (139.6 MeV) hace que dicho proceso no resulte todo lo óptimo que se requiere en la actual generación de análisis. Por otro lado, el detector SBND consta, además de la cámara de deriva que registra la ionización, de un sofisticado sistema de detección de luz que actúa de manera complementaria a éste. La reconstrucción de la señal de luz registrada en el experimento SBND se encuentra lo bastante madura como para que sea interesante explorar si dicho detector puede aportar vías de separación complementarias, que en este trabajo realizaremos mediante un análisis basado en técnicas de aprendizaje supervisado.

Tipología: Estudio de casos, teóricos o prácticos, relacionados con la temática del Grado.

Objetivos planteados:

Llevar a cabo una revisión bibliográfica sobre metodologías consolidadas de clasificación estadística, con especial referencia a su estado actual, y exposición sintética del conocimiento desde una perspectiva global en el contexto del análisis estadístico de datos multivariantes • Familiarizarse con la técnica de detección de las cámaras de deriva de argón líquido (LARTPCs), así como de los objetivos experimentales de SBND y del programa de corto recorrido (“short-baseline”) de búsqueda de neutrinos estériles • Profundizar en el análisis de datos según del modelo más

adecuado a un problema concreto, según el carácter de los datos multivariantes usados como entrada, con una clara identificación de los elementos conceptuales, y exposición de sus fundamentos matemáticos y aspectos metodológicos • Entender la estructura de los archivos de análisis de SBND • Realizar un estudio estadístico de viabilidad de separación en función de las distintas variables disponibles • Aplicar las técnicas arriba descritas a la separación de partículas en el estado final, mediante el uso y desarrollo eventual de procedimientos computacionales y gráficos relacionados • Escribir un informe final en base a los hallazgos, problemas encontrados y soluciones adoptadas

Bibliografía básica:

[1] C. C. Aggarwal, "Data Mining: The Textbook". Springer [2] F.E. Harrel, "Regression Modeling Strategies". Springer [3] B. Ratner, "Statistical and Machine-Learning Data Mining. Techniques for Better Predictive Modelling and Analysis of Big Data". CRC Press [4] D. Zeltermanm "Applied Multivariate Statistics with R". Springer [5] <http://sbn-nd.fnal.gov/> [6] P. Machado, O. Palamara, D. Schmitz, "The Short-Baseline Neutrino Program at Fermilab", Ann.Rev.Nucl.Part.Sci. 69 (2019) 363-387. <https://www.annualreviews.org/doi/10.1146/annurev-nucl-101917-020949> [7] <http://pdg.lbl.gov/2018/reviews/rpp2018-rev-neutrino-mixing.pdf> [8] M. Thomson, "Modern Particle Physics". Cambridge University Press

Recomendaciones y orientaciones para el estudiante:

En primer lugar, será necesario introducir los conceptos teóricos necesarios, así como las herramientas a utilizar. En particular, se llevará a cabo una identificación de los aspectos conceptuales inherentes a los enfoques de clasificación objeto de estudio y su formalización matemática. Asimismo, se realizará un análisis pormenorizado de los fundamentos matemáticos que sustentan el desarrollo de la metodología objeto central del trabajo. Fundamentalmente se utilizará R [4] para el estudio de los archivos de análisis de SBND. A continuación, se tratará de filtrar las variables de interés en base a algún criterio cualitativo y/o cuantitativo de separación de señal-fondo. Una vez que se haya demostrado que la separación es viable en función de cortes rígidos en estas variables, se procederá a utilizar soluciones de aprendizaje supervisado para optimizar el proceso. Por último, se desarrollará un estudio aplicado e interpretación de los resultados con referencia precisa a la fundamentación matemática de la metodología.

Plazas: 1

2. DATOS DEL TUTOR/A:

Nombre y apellidos: BRUNO ZAMORANO GARCÍA

Ámbito de conocimiento/Departamento: FÍSICA TEÓRICA

Correo electrónico: bzamorano@ugr.es

3. COTUTOR/A DE LA UGR (en su caso):

Nombre y apellidos: JOSÉ LUIS ROMERO BÉJAR

Ámbito de conocimiento/Departamento: ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

Correo electrónico: jlrbekar@ugr.es

4. COTUTOR/A EXTERNO/A (en su caso):

Nombre y apellidos:

Correo electrónico:

Nombre de la empresa o institución:

Dirección postal:

Puesto del tutor en la empresa o institución:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos: ALVARO MUÑOZ RUIZ

Correo electrónico: alvaromur2001@correo.ugr.es