UNIVERSIDAD DE GRANADA

FACULTAD DE CIENCIAS

Propuesta TFG. Curso 2026/2027

GRADO: Doble Grado en Física y Matemáticas

CÓDIGO DEL TFG: 295-030-2026/2027

1. DATOS BÁSICOS DEL TFG:

Título: Búsqueda de piones neutros en el experimento SBND mediante técnicas de "Machine Learning"

Descripción general (resumen y metodología):

El experimento SBND [1] [2] consiste en una cámara de deriva de argón líquido expuesta a un haz de neutrinos. Parte de su programa científico se centra en el estudio de la aparición de neutrinos del electrón a partir de oscilaciones de neutrinos del muon [3], a fin de revelar si las anomalías observadas en experimentos de oscilaciones de neutrinos de corto recorrido son compatibles con la existencia de un neutrino estéril. Alternativamente, una de las explicaciones alternativas a esta anomalía es una incorrecta estimación del fondo o background, en particular el correspondiente a sucesos con un pion neutro en el estado final. Esto es debido a que este mesón se desintegra casi exclusivamente mediante la interacción electromagnética, produciendo dos fotones en el estado final que, en determinadas circunstancias, pueden confundirse con la señal característica de un electrón.

Separar este tipo de sucesos de manera automática en base a observables experimentales es por tanto de gran interés para el análisis de oscilaciones. Además, la reconstrucción de la masa en reposo del pion neutro supone una excelente validación de la escala de energía del detector y permite evaluar el funcionamiento de la reconstrucción. La reconstrucción de los sucesos de neutrino en SBND es lo bastante madura como para que sea viable explorar técnicas de separación basadas en "Machine Learning". Estas técnicas se han empleado con éxito en otros experimentos de neutrinos, por lo que esperamos ser capaces de hacerlo también en SBND.

Los algoritmos que se utilizan en este tipo de aprendizaje suelen apoyarse en modelos probabilísticos, como la regresión lineal, la regresión logística, los árboles de decisión o las redes neuronales, entre otros. Para seleccionar el método de clasificación más adecuado, es fundamental considerar si las variables de entrada son de tipo cuantitativo o cualitativo, así como la distribución conjunta de probabilidad de dichas variables y otras propiedades relevantes del conjunto de datos. Según estas características, se podrá optar por la técnica de clasificación que mejor se adapte al problema. Cabe señalar que el aprendizaje supervisado no es la única estrategia de aprendizaje automático utilizada en disciplinas como la física. También existen otros enfoques, como el aprendizaje no supervisado, el aprendizaje por refuerzo y diversas técnicas de aprendizaje profundo. Cada uno de estos métodos presenta ventajas y limitaciones específicas, y su idoneidad dependerá del tipo de problema que se pretenda resolver.

En este trabajo se propone aplicar técnicas de aprendizaje supervisado para desarrollar un clasificador capaz de identificar sucesos de neutrinos del electrón en los datos simulados de SBND. Se explorarán distintos algoritmos y metodologías con el fin de maximizar la eficiencia de identificación y minimizar las tasas de falso positivo, lo que repercutirá directamente en la sensibilidad del experimento.

Metodología:

En primer lugar, será necesario introducir los conceptos teóricos necesarios, así como las herramientas a utilizar. Fundamentalmente se utilizará R o Python para el estudio de los archivos de análisis de SBND, así como herramienta preferida a la hora de aplicar las técnicas de "Machine Learning", por su sencillez de uso y modularidad. 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 “Machine Learning” para optimizar el proceso.

En esta fase se seguirán los procedimientos habitualmente prescritos en el campo, como son el uso de conjuntos de entrenamiento y prueba, “cross-validation” y optimización de hiperparámetros, aunque el principal énfasis se pondrá en la comprensión cualitativa de estas técnicas y en los criterios de selección que permiten decantarse por una u otra. Finalmente se evaluará el rendimiento del algoritmo escogido y se propondrán futuras líneas de mejora.

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

Objetivos planteados:

- Familiarizarse con la técnica de detección de las cámaras de deriva de argón líquido (LArTPCs)
- Entender los objetivos experimentales de SBND y del programa de corto recorrido (“short-baseline”) de búsqueda de neutrinos estériles
- 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
- Revisión bibliográfica exhaustiva sobre técnicas para la clasificación de datos multivariantes, identificando las principales metodologías y enfoques utilizados en la literatura, con especial referencia a su estado actual, y exposición sintética del conocimiento desde una perspectiva global.
- Introducción de 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
- Entrenar distintos algoritmos de “Machine Learning” y utilizar criterios genéricos para escoger el más adecuado en este caso. Como primera aproximación a estas técnicas se emplearán los algoritmos de kNN, “Boosted decision trees” y regresión penalizada. Si el tiempo lo permite se explorará asimismo el uso de redes neuronales sencillas (“Multilayer perceptron”).
- Escribir un informe final en base a los hallazgos, problemas encontrados y soluciones adoptadas

Bibliografía básica:

[1] - http://sbn-nd.fnal.gov/
[2] - 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
[3] - http://pdg.lbl.gov/2018/reviews/rpp2018-rev-neutrino-mixing.pdf
[4] - M. Thomson, “Modern Particle Physics”. Cambridge University Press
[5] - C. C. Aggarwal, “Data Mining: The Textbook”. Springer
[6] - B. Ratner, “Statistical and Machine-Learning Data Mining. Techniques for Better Predictive Modelling and Analysis of Big Data”. CRC Press

Recomendaciones y orientaciones para el estudiante:

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: jlrbejar@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:

Centro de convenio Externo:

5. DATOS DEL ESTUDIANTE:

Nombre y apellidos: NURIA TOURIÑO VERANO

Correo electrónico: ntouver@correo.ugr.es