



Propuesta de Trabajo Fin de Grado del Doble Grado en Física y Matemáticas (curso 2022-23)

Responsable de tutorización: Bruno Zamorano García

Correo electrónico: bzamorano@ugr.es

Departamento: Física Teórica y del Cosmos

Área de conocimiento: Física Teórica

Responsable de cotutorización: José Luis Romero Béjar

Correo electrónico: jlrbejar@ugr.es

Departamento: Estadística e Investigación Operativa

Área de conocimiento: Análisis Multivariante

(Rellenar solo en caso de que la propuesta sea de un estudiante):

Estudiante que propone el trabajo: Luis Hallamaa Carmona

Título: Técnicas de Aprendizaje Supervisado Aplicadas a clasificación de topologías de estado final en el experimento SBND

Número de créditos: 6 ECTS 12 ECTS

Tipología del trabajo (marcar las casillas que correspondan):

1. Revisiones y/o trabajos bibliográficos sobre el estado actual de aspectos específicos relacionados con la titulación

2. Estudio de casos, teóricos o prácticos, relacionados con la temática de la titulación, a partir de material disponible en los centros

3. Trabajos experimentales, de toma de datos de campo, de laboratorio, etc.

4. Elaboración de nuevas prácticas de laboratorio

5. Elaboración de un informe o un proyecto en el ámbito del grado de naturaleza profesional

6. Trabajos relacionados con las prácticas externas

Descripción y resumen de contenidos:

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.

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. 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, proporciona una información muy granular y detallada acerca de las partículas en el estado final (es decir, tras la interacción del neutrino con el núcleo). Por ello, se trata de un escenario idóneo en el que aplicar técnicas de aprendizaje supervisado.

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 topologías de estado final en el experimento SBND.

Actividades a desarrollar:

- 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.
- Revisión bibliográfica sobre 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.
- Profundización 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.
- Aplicación para la clasificación de topologías de estado final mediante el uso y desarrollo eventual de procedimientos computacionales y gráficos relacionados.

Objetivos planteados

Identificación de los aspectos conceptuales inherentes a los enfoques de clasificación objeto de estudio y su formalización matemática.

Análisis pormenorizado de los fundamentos matemáticos que sustentan el desarrollo de la metodología objeto central del trabajo.

Desarrollo de un estudio aplicado e interpretación de resultados con referencia precisa a la fundamentación matemática de la metodología.

Bibliografía

- [1] C. C. AGGARWAL, *Data Mining: The Textbook*, (Springer, Switzerland, 2015).
- [2] F. E. HARREL, *Regression Modeling Strategies*, (Springer, New York, 2001).
- [3] B. RATNER, *Statistical and Machine-Learning Data Mining. Techniques for Better Predictive Modelling and Analysis of Big Data*, (CRC Press, Boca Raton, 2011).
- [4] D. ZELTERMAN, *Applied Multivariate Statistics with R*, (Springer, Switzerland, 2015).
- [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.
- [7] M. THOMSON, *Modern Particle Physics*, (Cambridge University Press, 2015)

(Firmar solo en caso de trabajos propuestos por estudiantes)

Firma del estudiante

Firma del responsable de tutorización

Firma del responsable de cotutorización

En Granada, a 18 de mayo de 2022.