

Resumen Ejecutivo del Proyecto de Obtención de Grado

PREDICCIÓN DE EVENTOS EN LA HISTORIA DE VIDA UNIVERSITARIA CON DEEP LEARNING

Mtra. Gisel Hernández Chávez, ghernand@iteso.mx
Dr. José Francisco Cervantes Álvarez, fcervantes@iteso.mx

1) Descripción general del proyecto propuesto

Esta propuesta tiene como antecedente los modelos predictivos de eventos de historia de vida desarrollados por la Mtra. Hernández y los trabajos realizados en el campo de deep learning por el Dr. Cervantes. Actualmente contamos con modelos de estimación de abandono, cambio de carrera y egreso a nivel de la universidad y de grupos de alumnos (mujeres, hombres, por carreras, por rangos de edades, etc.), así como modelos predictivos a nivel individual. Para estos últimos hemos empleado técnicas estadísticas como regresión semi paramétrica de Cox y modelos de Cox extendidos y modelos de aprendizaje de máquina como los árboles de supervivencia aleatorios (RSF). Los modelos se han desarrollado con el fin principal de predecir la deserción en alumnos de licenciatura del ITESO.

Se empleará una metodología de Minería de Datos con elementos específicos del Análisis de Supervivencia propuesta por la Mtra. Hernández en su tesis doctoral.

El objetivo del trabajo es desarrollar modelos de supervivencia con Deep learning (Aastha et al., 2020; Borisov et al., 2021; Fotso, 2018; Grinsztajn et al., 2022; Gupta et al., 2019; Hao et al., 2021; Katzman et al., 2018; Lee et al., 2018, 2020; *Loss and Loss Functions for Training Deep Learning Neural Networks*, n.d.; *Survival Analysis for Deep Learning* | Sebastian Pölsterl, n.d.; Rietschel et al., 2018) y comparar la calidad de estos modelos con los resultados obtenidos con Cox regularizados (Li et al., 2017; Tay et al., 2022) y RSF (Afrin et al., 2018; Ishwaran et al., 2014; Ishwaran & Kogalur, 2010; Miao Fen and Cai, 2015; Mogensen et al., 2012; Taylor, 2011; Waititu et al., n.d.). Contamos con datos reales para el desarrollo del trabajo.

2) Objetivo General

Desarrollar modelos de supervivencia con Deep learning y comparar la calidad de estos, fundamentalmente en cuanto a precisión, con los resultados obtenidos con Cox regularizados y RSF.

3) Entregabe esperado

- Convenio de confidencialidad.
- TOG con Introducción, Estado del arte, Marco teórico, Metodología de minería de datos, Resultados y discusión, Conclusiones, Anexos y Bibliografía.
- Código y archivos generados que serán propiedad del ITESO.
- No es obligatorio, pero sí recomendable, la presentación en congreso y la escritura de un artículo científico.

4) Vinculación o colaboración

La vinculación fundamental es con asesores en el área de estadística del CUCEI (UdG) y el Dr. Jesús Eladio Sánchez García de la Academia de Ciencias de Cuba.

5) Asignaturas de la MSC relacionadas con el desarrollo del proyecto

Entre las materias electivas se recomienda tomar: Programación para Análisis de Datos, Análisis Estadístico Multivariado y cualquiera de Aprendizaje Profundo.

6) Participación en el proyecto

Puede participar más de un alumno, caso en el cual se dividirían con diferentes abordajes del tiempo (continuo o discreto), dado que en estos estudios longitudinales varían los modelos según esta asunción. Otro criterio de división del trabajo puede ser la ampliación a otras técnicas para comparar la precisión de modelos, como las máquinas de soporte vectorial (SVM) y otras técnicas basadas en árboles. Un tercer criterio de división es el abordaje desde un solo evento (por ejemplo, cambio de carrera) o eventos en competencia (simultáneamente estudiar varios eventos de historia de vida).

CV del proponente

Ingeniera en Sistemas Automatizados de Dirección 1981-1986, Instituto Superior Politécnico "José A. Echeverría" (ISPJAE), Cuba. Especialidad "Aseguramiento Matemático, Informativo y Técnico de los Sistemas Automatizados de Dirección", ISPJAE. Maestría "Sistemas de Información", CUCEI, UdG. Ingeniería en Sistemas Computacionales, 2003, ITESO, México. Doctorado en Tecnologías de Información, 2018-2022, CUCEA, UdG, en proceso de defensa de la tesis. Últimas publicaciones:

Hernández-Chávez, G., & López_Martin, C. (marzo de 2020). Caracterización del abandono según el género en la universidad ITESO. En VI Seminario SIEEE. Seminario llevado a cabo en San Luis Potosí, México.

Hernández Chávez, G., López Martín, C., & Sánchez García, J. E. (2020). Perfil de alumnos con abandono escolar en una universidad privada mexicana. CPU-e, Revista de Investigación Educativa, 0(31), 97-121. <https://doi.org/10.25009/cpue.v0i31.2703>

1986-1992	Profesora del ISPJAE, Facultad de Ingeniería Industrial, Cuba.
1992-1993	Especialista en Computación del Centro para la Automatización de la Industria Mecánica, Ministerio de la Industria Sidero-Mecánica, Cuba.
1993-2005	Profesora-Investigadora Titular A del Departamento de Ciencias Computacionales, CUCEI, Universidad de Guadalajara.
1995 -actual	Profesora del ITESO, Coord. de Ingeniería en Sistemas Computacionales (2000-2003), Coordinadora de la Incubadora de Empresas Tecnológicas (2004-2010). Miembro del Consejo de Rectoría (2006-217), Directora General Académica (2011-2017), Junta de Gobierno (2011-2017).
2015 a 2019	Program Advisory Council, School of Business and Technology, North Central University, USA.