

ASIGNATURA DE PROCESAMIENTO DEL LENGUAJE NATURAL

| | | | | | |
|--|--|-----------------|-------------------------|--------------|-----------------|
| PROPÓSITO DE APRENDIZAJE DE LA ASIGNATURA | Proporcionar a los estudiantes conocimientos sobre algoritmos, técnicas y enfoques metodológicos usados en el procesamiento del lenguaje natural, para su aplicación en el desarrollo de proyectos de investigación o desarrollo tecnológico. Al finalizar el curso, el alumno tendrá un panorama amplio sobre el tratamiento del lenguaje natural por técnicas automáticas. | | | | |
| CUATRIMESTRE | CUARTO | | | | |
| TOTAL DE HORAS | PRESENCIALES | NO PRESENCIALES | HORAS POR SEMANA | PRESENCIALES | NO PRESENCIALES |
| | 75 | 15 | | 5 | 1 |

| UNIDADES DE APRENDIZAJE | HORAS DEL SABER | | HORAS DEL SABER HACER | | HORAS TOTALES | |
|---|-----------------|----------|-----------------------|-----------|---------------|-----------|
| | P | NP | P | NP | P | NP |
| I. Fundamentos del procesamiento del lenguaje natural. | 10 | 0 | 15 | 5 | 25 | 5 |
| II. Modelos tradicionales del lenguaje. | 10 | 0 | 15 | 5 | 25 | 5 |
| III. Aprendizaje automático para el procesamiento del lenguaje natural. | 10 | 0 | 15 | 5 | 25 | 5 |
| TOTALES | 30 | 0 | 45 | 15 | 75 | 15 |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

COMPETENCIA A LA QUE CONTRIBUYE LA ASIGNATURA

De acuerdo con la metodología de diseño curricular de la DGUTyP, las competencias se desagregan en dos niveles de desempeño: Unidades de Competencias y Capacidades.

La presente asignatura contribuye al logro de la competencia y los niveles de desagregación descritos a continuación:

COMPETENCIA: Conocer, analizar y aplicar conocimientos, así como técnicas de procesamiento de lenguaje natural (PLN) para el desarrollo de investigación y desarrollo tecnológico en una amplia variedad de áreas.

| UNIDADES DE COMPETENCIA | CAPACIDADES | CRITERIOS DE DESEMPEÑO |
|---|---|---|
| Comprender los fundamentos teóricos y prácticos para la aplicación de técnicas de PLN a la solución de problemas multidisciplinares que requieren interacción en lenguajes naturales. | Conocer los niveles de análisis del lenguaje y las técnicas más comunes para tratar los desafíos lingüísticos relacionados a cada nivel | Solución de casos de estudio referentes a la exploración de conjuntos de datos y técnicas de preprocesamiento <ul style="list-style-type: none"> - Análisis de la tarea a resolver - Exploración del conjunto los conjuntos de datos - Preprocesamiento - Análisis del uso de recursos léxicos |
| | Conocer los recursos léxicos disponibles en el marco del Procesamiento del Lenguaje Natural. | |
| | Conocer e implementar modelos del lenguaje para transformar el lenguaje natural, en sus diferentes modalidades, a representaciones que alimenten algoritmos de aprendizaje automático | Solución de casos de estudio referentes a la construcción de modelos para solucionar problemas que requieren procesamiento del lenguaje natural <ul style="list-style-type: none"> - Análisis de la tarea a resolver - Representación de los datos - Preprocesamiento e ingeniería de características - Generación del modelo del lenguaje Análisis del modelo del lenguaje |
| | Conocer y aplicar técnicas de aprendizaje automático y profundo para crear modelos predictivos que tomen como entrada las representaciones creadas | Solución de casos de estudio referentes a la aplicación de redes neuronales: <ul style="list-style-type: none"> - Aplicación de distintas arquitecturas de redes neuronales profundas |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

| | | |
|--|--|---|
| | analizando contenido en lenguaje natural | <ul style="list-style-type: none"> - Parametrización de los modelos - Evaluación de los modelos Aplicación de técnicas de transferencia de aprendizaje en PLN |
|--|--|---|

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

UNIDADES DE APRENDIZAJE

| | | | | | | | | |
|------------------------------|--|-----------|------------------------|----------|-----------|------------------------------|----------|-----------|
| UNIDAD DE APRENDIZAJE | I. Fundamentos del procesamiento del lenguaje natural. | | | | | | | |
| PROPÓSITO ESPERADO | El alumno conocerá y aplicará los algoritmos, técnicas y enfoques metodológicos para el tratamiento automático de los fenómenos léxicos, sintácticos y semánticos del lenguaje humano principales técnicas y enfoques metodológicos usados en el procesamiento | | | | | | | |
| HORAS TOTALES | P | NP | HORAS DEL SABER | P | NP | HORAS DEL SABER HACER | P | NP |
| | 25 | 5 | | 10 | 0 | | 15 | 5 |

| TEMAS | SABER DIMENSIÓN CONCEPTUAL | SABER HACER DIMENSIÓN ACTUACIONAL | SER DIMENSIÓN SOCIOAFECTIVA |
|---|--|--|---|
| Niveles de análisis del lenguaje y | <p>Explicar los conceptos generales de los sistemas de interacción en lenguajes naturales.</p> <p>Conocer los recursos léxicos y conjuntos de datos (corpora) más comunes para el diseño de herramientas basadas en tratamientos estadísticos y de aprendizaje automático</p> <p>Identificar las técnicas de preprocesamiento más comunes de acuerdo con la modalidad de los datos a procesar.</p> | <p>Investigar los principales niveles de lenguaje y sus principales técnicas de análisis.</p> <p>Conocer diferentes recursos léxicos y las técnicas de creación de los conjuntos de datos</p> <p>Aplicar técnicas de preprocesamiento de acuerdo con la modalidad de los datos a procesar.</p> | <p>Analítico</p> <p>Proactivo</p> <p>Autónomo</p> <p>Reflexivo</p> <p>Responsable</p> <p>Ordenado</p> <p>Observador</p> <p>Disciplinado</p> |
| Tareas del Procesamiento del Lenguaje Natural | Exponer las principales tareas del Procesamiento del Lenguaje Natural | Investigar los fundamentos, principales desafíos y tendencias en las tareas del Procesamiento del Lenguaje Natural | <p>Analítico</p> <p>Proactivo</p> <p>Autónomo</p> <p>Responsable</p> <p>Ordenado</p> <p>Observador</p> <p>Disciplinado</p> |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

| PROCESO DE EVALUACIÓN | | TÉCNICAS SUGERIDAS DE ENSEÑANZA Y APRENDIZAJE | ESPACIO DE FORMACIÓN | | | MATERIALES Y EQUIPOS |
|---|---|---|----------------------|------------|----------|---|
| EVIDENCIA DE DESEMPEÑO | INSTRUMENTO EVALUACIÓN | | AU LA | TAL LER | OT RO | |
| Exploración y preprocesamiento de conjuntos de datos para resolver problemas relacionados con las tareas del Procesamiento del Lenguaje Natural | Reportes de prácticas de laboratorio. Ejercicios prácticos. Proyecto. | Solución de problemas Exposición Tareas de investigación Prácticas Diseño de sistemas de clasificación supervisada usando software especializado. | X | X | | Equipo de cómputo. Pizarrón. Plumón. Material impreso. Software especializado. Internet. |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

| | | | | | | | | |
|------------------------------|---|-----------|------------------------|----------|-----------|------------------------------|----------|-----------|
| UNIDAD DE APRENDIZAJE | II. Modelos tradicionales del lenguaje. | | | | | | | |
| PROPÓSITO ESPERADO | El alumno conocerá los modelos tradicionales del lenguaje para su aplicación en la solución de problemas del área del Procesamiento del Lenguaje Natural. | | | | | | | |
| HORAS TOTALES | P | NP | HORAS DEL SABER | P | NP | HORAS DEL SABER HACER | P | NP |
| | 25 | 5 | | 10 | 0 | | 15 | 5 |

| TEMAS | SABER DIMENSIÓN CONCEPTUAL | SABER HACER DIMENSIÓN ACTUACIONAL | SER DIMENSIÓN SOCIOAFECTIVA |
|---|---|---|---|
| Modelos estadísticos del lenguaje | Introducir los tipos de características más comunes para problemas de PLN Explicar los conceptos fundamentales de los modelos del lenguaje más comunes en PLN. | Identificar y extraer las características de los diferentes modelos del lenguaje. Generar diferentes modelos del lenguaje de acuerdo con la modalidad de los datos a procesar. | Analítico Proactivo Autónomo Responsable Ordenado Observador Disciplinado |
| Incrustaciones de palabras (embeddings) | Introducir los conceptos fundamentales de las incrustaciones de palabras | Construir representaciones basadas en incrustaciones de palabras | Analítico Proactivo Autónomo Responsable Ordenado Observador Disciplinado |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

| PROCESO DE EVALUACIÓN | | TÉCNICAS SUGERIDAS DE ENSEÑANZA Y APRENDIZAJE | ESPACIO DE FORMACIÓN | | | MATERIALES Y EQUIPOS |
|--|---|---|-------------------------|------------|----------|--|
| EVIDENCIA DE DESEMPEÑO | INSTRUMENTO EVALUACIÓN | | AU LA | TAL LER | OT RO | |
| Implementación de modelos tradicionales del lenguaje usando conjuntos de datos disponibles | Reportes de prácticas de laboratorio. Ejercicios prácticos. Proyecto. | Solución de problemas Diseño de sistemas ópticos usando software especializado. Exposición Tareas de investigación Prácticas de laboratorio | X | X | | Material y equipo de laboratorio. Pizarrón. Plumón. Material impreso. Software especializado. Computadora. Internet. |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

| | | | | | | | | |
|------------------------------|---|-----------|------------------------|----------|-----------|------------------------------|----------|-----------|
| UNIDAD DE APRENDIZAJE | III. Aprendizaje automático para el procesamiento del lenguaje natural. | | | | | | | |
| PROPÓSITO ESPERADO | El alumno integrará conocimientos de Procesamiento del Lenguaje Natural y Aprendizaje Automático para generar enfoques de solución a problemas que requieren el tratamiento de lenguajes naturales. | | | | | | | |
| HORAS TOTALES | P | NP | HORAS DEL SABER | P | NP | HORAS DEL SABER HACER | P | NP |
| | 25 | 5 | | 10 | 0 | | 15 | 5 |

| TEMAS | SABER DIMENSIÓN CONCEPTUAL | SABER HACER DIMENSIÓN ACTUACIONAL | SER DIMENSIÓN SOCIOAFECTIVA |
|---|--|--|---|
| Aprendizaje automático para el Procesamiento del Lenguaje Natural | Introducir el marco conceptual general del aprendizaje automático aplicado en la solución de problemas del aprendizaje automático. | Aplicar diferentes algoritmos de aprendizaje automático para resolver tareas de PLN usando diferentes modalidades del lenguaje | Analítico Proactivo Autónomo Responsable Ordenado Observador Disciplinado |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

| PROCESO DE EVALUACIÓN | | TÉCNICAS SUGERIDAS DE ENSEÑANZA Y APRENDIZAJE | ESPACIO DE FORMACIÓN | | | MATERIALES Y EQUIPOS |
|--|---|---|-------------------------|------------|----------|--|
| EVIDENCIA DE DESEMPEÑO | INSTRUMENTO EVALUACIÓN | | AU LA | TAL LER | OT RO | |
| Implementación de técnicas de aprendizaje automático para el tratamiento computacional del lenguaje natural. | Reportes de prácticas de laboratorio. Ejercicios prácticos. Proyecto. | Solución de problemas Diseño de sistemas ópticos usando software especializado. Exposición Tareas de investigación Prácticas de laboratorio | X | X | | Material y equipo de laboratorio. Pizarrón. Plumón. Material impreso. Software especializado. Computadora. Internet. |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |

REFERENCIAS BIBLIOGRÁFICAS

| AUTOR | AÑO | TÍTULO DEL DOCUMENTO | LUGAR DE PUBLICACIÓN | EDITORIAL | ISBN |
|---|------|--|-----------------------|-----------------------------------|-------------------|
| <i>Zhang Yue, Teng Zhiyang</i> | 2021 | <i>Natural Language Processing</i> | <i>Reino Unido</i> | <i>Cambridge University Press</i> | 978-3-030-26621-9 |
| <i>Kevin P. Murphy</i> | 2012 | <i>Machine Learning: A Probabilistic Perspective</i> | <i>Estados Unidos</i> | <i>MIT Press</i> | 0262018020 |
| <i>Daniel Jurafsky, James H. Martin</i> | 2008 | <i>Speech and Language Processing</i> | <i>Estados Unidos</i> | <i>Prentice Hall</i> | 9780131873216 |
| <i>Ian Goodfellow, Yoshua Bengio, Aaron Courville</i> | 2016 | <i>Deep Learning</i> | <i>Estados Unidos</i> | <i>MIT Press</i> | 0262035618 |
| <i>Chollet, Francois</i> | 2017 | <i>Deep Learning with Python</i> | <i>Estados Unidos</i> | <i>Manning Publications</i> | 1617294438 |
| | | | | | |
| | | | | | |

| | | | |
|-----------------|---|-----------------------------------|---|
| ELABORÓ: | Comité del Doctorado en Optomecatrónica de la UPT | REVISÓ: | Dirección de Investigación y Posgrado de la UPT |
| APROBÓ: | DGUTyP | FECHA DE ENTRADA EN VIGOR: | Enero 2022 |