

## SYLLABUS: PLN, LLMs y Agentes de IA

### Instructor(es):

- Francisco Vidal
- Manuel Bozo
- Vicente Wohl (Laboratorios)

### 1. Introducción

Este curso ofrece un recorrido intensivo y completo, desde los fundamentos del Procesamiento de Lenguaje Natural<sup>1</sup> hasta la frontera actual de la IA. Iniciaremos construyendo una base sólida con modelos clásicos y redes recurrentes, para luego sumergirnos en la era moderna dominando la arquitectura Transformer y los LLMs<sup>2</sup> como BERT<sup>3</sup> y GPT<sup>4</sup>. Analizando códigos de soluciones ya implementadas, comprenderán cómo se aplican y ajustan estos modelos para resolver tareas complejas como clasificación, resumen y respuesta a preguntas.

La etapa final del curso los llevará a la vanguardia del campo, donde aprenderán a diseñar e implementar agentes inteligentes que utilizan LLMs como su núcleo de razonamiento. Exploraremos cómo dotar a estos sistemas de memoria, acceso a herramientas externas y la capacidad de planificar y actuar en entornos complejos mediante técnicas como RAG<sup>5</sup>. Al concluir, estarán capacitados no solo para usar modelos, sino para ser arquitectos de sistemas de IA autónomos y de extremo a extremo.

### 2. Objetivos

Al termino de esta capacitación, el participante será capaz de:

- Explicar y aplicar los principios computacionales y lingüísticos que subyacen al procesamiento del lenguaje natural, diferenciando entre representaciones simbólicas, estadísticas y neuronales (Word2Vec<sup>6</sup>, RNNs<sup>7</sup>, LSTMs<sup>8</sup>).
- Describir en detalle el funcionamiento de la arquitectura Transformer y el mecanismo de atención, y aplicar este conocimiento para utilizar y adaptar (fine-tuning) LLMs como BERT y GPT en tareas complejas (clasificación, QA<sup>9</sup>, NER<sup>10</sup>).
- Construir sistemas autónomos que utilicen LLMs como núcleo de razonamiento, integrando componentes clave como memoria (vectorial y conversacional), acceso a herramientas externas y patrones de

planificación (ReAct<sup>11</sup>, Plan-and-Execute).

- Desarrollar y evaluar pipelines de RAG de extremo a extremo, desde la creación de una base de conocimiento vectorial hasta la generación de respuestas contextualizadas y precisas.

### 3. Metodología

El curso se desarrolla en un formato de e-learning con sesiones en vivo los martes y viernes a mediodía, donde el instructor enseñará los conceptos clave de IA, PLN y LLMs. La aplicación práctica se demostrará a través del análisis de ejemplos y soluciones de IA ya construidas, explicando su lógica y arquitectura. Todas las clases son grabadas y quedan disponibles de forma permanente en la plataforma junto a los materiales de apoyo, permitiendo a los estudiantes aprender a su propio ritmo y usar estos casos como referencia para sus propios desarrollos.

### 4. Bibliografía y Recursos

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.

Bommasani, R., et al. (2021). *On the Opportunities and Risks of Foundation Models*. Center for Research on Foundation Models (CRFM), Stanford Institute for Human-Centered AI (HAI). arXiv preprint arXiv:2108.07258.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).

Hochreiter, S., & Schmidhuber, J. (1997). *Long Short-Term Memory*. Neural Computation, 9(8), 1735–1780.

[HuggingFace: Other AI Learning Courses](#)

[HuggingFace: Agents Course](#)

[HuggingFace: Deep Reinforcement Learning Course](#)

[HuggingFace: LLM Course](#)

[HuggingFace: Model Context \(MCP\) Protocol Course](#)

Jurafsky, D., & Martin, J. H. *Speech and Language Processing*. (3rd ed. draft). Prentice Hall.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. arXiv preprint arXiv:1907.11692.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. arXiv preprint arXiv:1301.3781.

Ouyang, L., et al. (2022). *Training language models to follow instructions with human feedback*. Advances in Neural Information Processing Systems 35.

Pennington, J., Socher, R., & Manning, C. D. (2014). *GloVe: Global Vectors for Word Representation*. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer*. Journal of Machine Learning Research, 21(140), 1-67.

Shinn, N., Cassano, F., Gopinath, A., & Yao, S. (2023). *Reflexion: Language Agents with Verbal Reinforcement Learning*. arXiv preprint arXiv:2303.11366.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. Advances in Neural Information Processing Systems 30 (NIPS 2017).

Yao, S., Zhao, J., Yu, D., Du, N., Tsvetkov, Y., & LeCun, Y. (2022). *ReAct: Synergizing Reasoning and Acting in Language Models*. arXiv preprint arXiv:2210.03629.

## 5. Trayectoria profesional Instructor(es)

Instructor: Francisco Vidal Ureta

- Ingeniero Comercial, Universidad de Valparaíso.
- Diplomado en Finanzas, Universidad de Chile.
- Diplomado en Big Data & Data Science. Pontificia Universidad Católica de Valparaíso.
- Applied Data Science Program, University of Michigan.
- Magister en Economía, Universidad de Chile.
- Magister en Inteligencia Artificial, Universidad Adolfo Ibáñez.

Instructor: Manuel Bozo Sáez

- Ingeniero en Información y Control de Gestión, Universidad de Chile.
- Professional Education, Massachusetts Institute of Technology.
- Exponential Learning, IE Business School.
- UN CC: e-Learn, United Nations Institute for Training and Research.

Instructor: Vicente Wohl Valderrama

- Ingeniero Civil Industrial con mención Financiera, Universidad de los Andes.
- Diplomado en Ciencia de Datos para la Gestión, Pontificia Universidad Católica de Chile
- Certificado profesional de Machine Learning Engineer, Google Cloud.

## 6. Programa

### **NIVEL 1: Fundamentos de PLN y Modelos Clásicos (DL<sup>12</sup> y texto)**

#### **SECCIÓN 1 (2 Clases)**

Tema: Fundamentos del Lenguaje (Avanzado): De Texto a Vectores

Objetivo: El participante debería comprender los siguientes puntos:

- Analizar los fundamentos lingüísticos (morfología, sintaxis, semántica) y las técnicas de pre-procesamiento de texto asociadas (tokenización, lematización, stemming).
- Diferenciar entre las representaciones de texto clásicas basadas en frecuencia, como BoW<sup>13</sup> y TF-IDF<sup>14</sup>, entendiendo sus ventajas y limitaciones.
- Comprender el concepto de word embeddings como vectores densos que capturan significado semántico, marcando el paso de la estadística al DL.
- Familiarizarse con la arquitectura y la intuición detrás de Word2Vec como modelo fundamental para la representación vectorial del lenguaje.
- Conocer técnicas de visualización como t-SNE<sup>15</sup> para interpretar las relaciones semánticas en un espacio vectorial de embeddings.

#### **SECCIÓN 2 (2 Clases + Laboratorio)**

Tema: Modelos Clásicos y Memoria: Procesamiento de Secuencias

Objetivo: Al finalizar la sesión, el participante será capaz de:

- Analizar el rol del modelado secuencial en PLN y el mecanismo operativo de modelos estadísticos clásicos, como los n-gramas y los HMMs<sup>16</sup>.
- Adquirir los conocimientos sobre la arquitectura fundamental de las RNNs y su mecanismo de "memoria" a través del estado oculto.
- Identificar el problema del desvanecimiento del gradiente en las RNNs y por qué limita su capacidad para aprender dependencias a largo plazo.
- Comprender la arquitectura de las redes LSTM y GRU<sup>17</sup>, con un enfoque en sus mecanismos de compuertas (gates) para gestionar el flujo de información.
- Familiarizarse con la aplicación de estos modelos secuenciales en tareas núcleo de PLN como la clasificación de texto, el análisis de sentimiento, modelado de temas (aplicando el modelo de análisis discriminante lineal) y el NER.

## **NIVEL 2: Modelos de Lenguaje Modernos y Aplicaciones Avanzadas**

### **SECCIÓN 3 (3 Clases)**

Tema: La Revolución Transformer: Atención y Arquitectura

Objetivo: Al término de este módulo, el participante habrá asimilado los siguientes conceptos:

- Analizar las limitaciones inherentes a las arquitecturas recurrentes (RNN/LSTM) que impulsaron el desarrollo de los Transformers.
- Adquirir un conocimiento profundo del mecanismo de auto-atención (self-attention), el componente clave del paper "Attention Is All You Need".
- Familiarizarse con la arquitectura completa de un Transformer, incluyendo los bloques Encoder y Decoder, la atención multi-cabeza (Multi-Head Attention) y las redes Feed-Forward.
- Entender cómo los Transformers manejan el orden de las secuencias sin recurrencia a través de los encodings posicionales.
- Comprender el concepto de embeddings contextualizados, y por qué representan un avance fundamental sobre embeddings estáticos como Word2Vec.

### **SECCIÓN 4 (2 clases + 1 Laboratorio)**

Tema: El Ecosistema de los Modelos: Pre-entrenamiento, Fine-Tuning y Aplicaciones

Objetivo: Para esta sesión, se espera que el participante logre:

- Comprender el paradigma de pre-entrenamiento y fine-tuning (transferencia de aprendizaje), la estrategia dominante para aplicar LLMs.
- Familiarizarse con las familias de modelos pre-entrenados más influyentes y sus diferencias arquitectónicas: BERT (encoder-only), GPT (decoder-only) y T5 (encoder-decoder).

- Conocer el rol del ecosistema Hugging Face como la librería y el hub central para acceder, utilizar y compartir modelos de lenguaje.
- Aprender cómo aplicar la técnica de fine-tuning para adaptar estos modelos a tareas núcleo de PLN: clasificación, QA, NER y summarización.
- Identificar las métricas de evaluación clave para tareas generativas y de clasificación: F1, BLEU<sup>18</sup>, ROUGE<sup>19</sup> y BERTScore.

### **NIVEL 3: Agentes Inteligentes Basados en LLMs**

#### **SECCIÓN 5 (2 Clases + Laboratorio)**

Tema: La Emergencia de los Agentes de IA: Razonamiento, Memoria y Herramientas

Objetivo: Los puntos clave a dominar en esta sesión incluyen:

- Comprender la transición de un LLM a un Agente de IA, y el rol del Prompt Engineering avanzado (Few-shot, Chain-of-Thought) para desbloquear capacidades de razonamiento.
- Familiarizarse con los patrones de razonamiento y acción fundamentales para agentes, como ReAct y Plan-and-Execute.
- Conocer los principales frameworks para la construcción de agentes, como LangChain, LlamaIndex y DSPy<sup>20</sup>, y entender su arquitectura modular (modelos, prompts, herramientas, memoria).
- Analizar la arquitectura RAG, explicando cómo su mecanismo de recuperación y generación sirve para fundamentar las respuestas de los agentes en conocimiento externo y mitigar alucinaciones.
- Adquirir los conocimientos sobre el uso de bases de datos vectoriales como la memoria a largo plazo de un agente, permitiendo la recuperación semántica de información.

#### **SECCIÓN 6 (2 Clases + Laboratorio)**

Tema: Sistemas y Alineamiento: Optimización, Colaboración y Seguridad de Agentes

Objetivo: Los objetivos de aprendizaje para esta sesión son:

- Aprender técnicas de optimización para adaptar y ejecutar agentes de manera eficiente, como LoRA<sup>21</sup> para el fine-tuning y estrategias de caching para reducir la latencia y los costos.
- Familiarizarse con arquitecturas multi-agente, incluyendo patrones de comunicación como MCP<sup>22</sup> y estructuras jerárquicas (líder/seguidor) para resolver problemas complejos de forma colaborativa.
- Comprender el concepto de alineamiento y la técnica de RLHF<sup>23</sup> como el principal mecanismo para guiar el comportamiento de los modelos hacia objetivos humanos.

- Adquirir los conocimientos para evaluar y auditar agentes, identificando sesgos, comportamientos tóxicos y fallos de seguridad.
- Analizar los principios de la IA responsable y los riesgos de seguridad en agentes, para aplicar marcos éticos en el diseño y despliegue de sistemas autónomos.

## 7. Glosario de abreviaciones

1. PLN: Procesamiento de Lenguaje Natural (*Natural Language Processing*).
2. LLM: Modelos de Gran Lenguaje (*Large Language Model*).
3. BERT: Representación de Codificador Bidireccional de Transformadores (*Bidirectional Encoder Representations from Transformers*).
4. GPT: Transformador Generativo Preentrenado (*Generative Pre-Trained Transformer*).
5. RAG: Generación Aumentada por Recuperación (*Retrieval-Augmented Generation*).
6. Word2Vec: Palabra a Vector (representación vectorial de palabras).
7. RNN: Redes Neuronales Recurrentes.
8. LSTM: Red (Neuronal) de Memoria a Corto-Largo Plazo (*Long Short-Term Memory*).
9. QA: Pregunta-Respuesta (*Question-Answering*).
10. NER: Reconocimiento de Entidades Nombradas (*Named-Entity Recognition*).
11. ReAct: Razonamiento y Acción (*Reasoning+Action*).
12. DL: Aprendizaje Profundo (*Deep Learning*).
13. BoW: Bolsa de Palabras (*Bag-of-Words*).
14. TF-IDF: Frecuencia de Término - Frecuencia Inversa de Documento (*Term Frequency-Inverse Document Frequency*).
15. t-SNE: T -Distribuido Estocástico Vecino Incrustación (*t-distributed Stochastic Neighbor Embedding*).
16. HMM: Modelos Ocultos de Márkov.
17. GRU: Unidad Recurrente Cerrada (*Gated Recurrent Units*).
18. BLEU: Evaluación Bilingüe Suplente (*Bilingual Evaluation Understudy*).
19. ROUGE: Estudio Orientado a la Memoria para la Evaluación de la Estructura Fundamental (*Recall-Oriented Understudy for Gisting Evaluation*).
20. DSPy: Marco de Secuenciación Declarativa de Python (*Declarative Sequencing Python Framework*).
21. LoRA: Adaptación de Bajo Rango (*Low-Rank Adaptation*).

22. MCP: Protocolo de Contexto de Modelo (*Model Context Protocol*).
23. RLHF: Aprendizaje por Refuerzo a Partir de la Retroalimentación Humana (*Reinforcement Learning from Human Feedback*).