

Índice

| | |
|---|-----------|
| MÓDULO 1. Procesamiento de texto en el PC | 2 |
| Importancia de los vectores en el aprendizaje automático (ML) y el análisis de datos | 2 |
| Definición de un vector | 2 |
| Utilidades específicas de los vectores en NLP | 3 |
| Desafíos y consideraciones al convertir texto en vectores | 3 |
| Bolsa de palabras (BoW – Bag of Words) | 3 |
| Enfoque del NLP | 3 |
| Representación de las palabras | 4 |
| Aplicaciones de BoW | 4 |
| Valor y eficacia de BoW | 4 |
| Conteo de palabras | 4 |
| Documento | 4 |
| Método de conteo | 5 |
| Tokenización | 5 |
| Maapeo | 5 |
| Tokenización | 5 |
| ejemplo 1 | 5 |
| Diferentes perspectivas para abordar el tema de tokenización | 6 |
| Consideraciones | 6 |
| Volumen de datos | 6 |
| Manejo de casos | 6 |
| ejemplo 2 | 6 |
| Stop Words | 7 |
| ejemplo 3 | 7 |
| ejemplo 4 | 8 |
| Stemming y Lemmatización | 8 |
| Stemming | 8 |
| Lemmatization | 8 |
| Aplicación de Stemming y Lemmatization en situaciones reales | 8 |
| ejemplo 5 | 9 |
| ejemplo 6 | 10 |
| Ejemplo práctico: Proyecto 1 (ver Anexo 1) | 10 |
| Similitud de vectores | 11 |
| Cálculo de similitud de vectores | 11 |
| Comparación del ángulo entre vectores y la distancia euclidiana | 12 |
| Método TF-IDF | 12 |
| Ambigüedad y especificidad de las stop words dependiendo de la aplicación | 12 |
| Función del TF-IDF | 13 |
| Fórmulas de TF-IDF | 13 |

| | |
|---|---------------|
| Aplicación del método TF-IDF ----- | 14 |
| Ejemplo práctico: Proyecto 2 (ver Anexo 2) ----- | 14 |
| Neural Word Embeddings ----- | 14 |
| Modelos para secuencias de Deep Learning ----- | 14 |
| Word Embeddings (Incrustación de palabras) ----- | 15 |
| Uso práctico de Word Embeddings ----- | 16 |
| Aplicación del Word Embeddings – Analogías ----- | 16 |
| Ejemplo práctico: Proyecto 3 (ver Anexo 3) ----- | 16 |
| Creación de un Embedding con Word2Vec ----- | 16 |
| ejemplo 7: CPU disponible en mi PC ----- | 17 |
| Ejemplos prácticos: Proyecto 4 (ver Anexo 4) ----- | 17 |
| Ejemplos prácticos: Proyecto 5 (ver Anexo 5) ----- | 17 |
| Ejemplos prácticos: Proyecto 6 (ver Anexo 6) ----- | 17 |
| Ejemplos prácticos: Proyecto 7 (ver Anexo 7) ----- | 17 |
| MÓDULO 2. Visión probabilística en AI ----- | 18 |
| Introducción a los modelos de Markov en NLP ----- | 18 |
| Aplicabilidad universal de los modelos de Markov ----- | 18 |
| Propiedad fundamental de los modelos de Markov ----- | 18 |
| Estructura y entrenamiento de un modelo de Markov ----- | 18 |
| Aplicaciones en NLP ----- | 19 |
| Procesos de Markov ----- | 19 |
| Representación matemática ----- | 20 |
| Uso de modelos de Markov para secuencias ----- | 20 |
| Estados en modelos de Markov ----- | 20 |
| Transición de estados y matrices de transición ----- | 21 |
| Implementación computacional y entrenamiento ----- | 22 |
| Procesos de Markov – Suavizado de Probabilidades ----- | 22 |
| Estimaciones de máxima verosimilitud ----- | 22 |
| Problemas de valores cero ----- | 22 |
| Suavizado de probabilidad ----- | 23 |
| Probabilidad de una secuencia ----- | 23 |
| Espacio logarítmico ----- | 24 |
| Construcción de un clasificador de texto ----- | 25 |
| Otras aplicaciones de un clasificador de texto ----- | 26 |
| Aplicación del modelo de Markov y clasificación de textos ----- | 26 |
| Regla de Bayes ----- | 26 |
| Aplicación de clasificador de textos ----- | 27 |
| Ejemplo práctico: Proyecto 8 (ver Anexo 8) ----- | 27 |
| Generador de texto ----- | 27 |
| Ampliación de los modelos de Markov ----- | 27 |
| Modelo de Markov de segundo orden ----- | 28 |

| | |
|--|-----------|
| Implicaciones----- | 28 |
| Creación de un generador de texto con Python ----- | 29 |
| Ejemplo práctico: Proyecto 9 (ver Anexo 9) ----- | 29 |
| Text Spinning ----- | 29 |
| Importancia de los motores de búsqueda ----- | 29 |
| Formas de hacer spinning de contenido ----- | 29 |
| Evolución de las técnicas de generación de contenido ----- | 29 |
| Avances técnicos en NLP----- | 30 |
| N-Gram ----- | 30 |
| Propuesta de modelo para Spinning ----- | 31 |
| Creación de Spinning de texto con python ----- | 31 |
| Ejemplo práctico: Proyecto 10 (ver Anexo 10) ----- | 31 |
| MÓDULO 3. Métodos de Machine Learning para NLP ----- | 32 |
| Introducción al aprendizaje automático----- | 32 |
| Ejemplos de modelos de aprendizaje ----- | 32 |
| Detección de Spam ----- | 32 |
| Importancia de la detección de Spam ----- | 33 |
| Automatización mediante ML para filtrar Spam ----- | 33 |
| Descripción del proceso----- | 33 |
| Regla de Naive Bayes ----- | 33 |
| Aplicación de la regla de Bayes en ML ----- | 35 |
| Aplicando Naive Bayes en ML ----- | 35 |
| Elección del modelo----- | 35 |
| Gaussiano ----- | 36 |
| Multinomial----- | 36 |
| Bernoulli ----- | 36 |
| Aplicación de detector de Spam----- | 37 |
| Ejemplo práctico: Proyecto 11 (ver Anexo 11) ----- | 37 |
| Análisis de Sentimientos ----- | 37 |
| Clasificación vs Regresión, en el análisis de sentimientos ----- | 38 |
| Descripción de la tarea del Análisis de Sentimiento ----- | 38 |
| Aplicabilidad y beneficios del Análisis de Sentimiento ----- | 38 |
| Regresión Logística ----- | 38 |
| Perspectiva vectorial sobre la tarea de clasificación ----- | 39 |
| Representación lineal ----- | 39 |
| Activación ----- | 40 |
| Funciones de activación----- | 41 |
| Función Sigmoide----- | 41 |
| Regresión logística Multiclases MLR ----- | 41 |
| Función SoftMax ----- | 41 |

| | |
|---|-----------|
| Salidas esperadas de la regresión logística multiclase----- | 42 |
| Aplicación de Análisis de Sentimiento ----- | 42 |
| Ejemplo práctico: Proyecto 12 (ver Anexo 12)----- | 42 |
| Ejemplo práctico: Proyecto 13 (ver Anexo 13)----- | 43 |
| Resumen de texto ----- | 43 |
| Sumarización con Inteligencia Artificial----- | 43 |
| Aplicaciones en motores de búsqueda----- | 43 |
| Categorización de la sumarización de texto----- | 44 |
| Proceso de sumarización----- | 44 |
| Resumen de texto con vectores ----- | 44 |
| Ventajas----- | 44 |
| Pasos básicos en la técnica TF-IDF para sumarización----- | 45 |
| Selección de sentencias para hacer el resumen----- | 46 |
| Consideraciones de cada método----- | 46 |
| Aplicación de resumen de texto con vectores ----- | 46 |
| Ejemplo práctico: Proyecto 14 (ver Anexo 14)----- | 46 |
| Resumen con TextRank ----- | 47 |
| Comparación de TextRank con TF-IDF----- | 47 |
| ¿Cómo funciona PageRank de Google?----- | 47 |
| Aplicación de TextRank----- | 48 |
| Proceso del TextRank----- | 48 |
| Paso a paso del TextRank----- | 49 |
| Aplicación de resumen de textos utilizando TextRank ----- | 49 |
| Ejemplo práctico: Proyecto 15 (ver Anexo 15)----- | 49 |
| Modelación de temas con LDA----- | 50 |
| Aplicaciones del LDA----- | 50 |
| Aprendizaje No supervisado con LDA----- | 50 |
| Outputs and Inputs of LDA----- | 51 |
| Método de conteo----- | 51 |
| LDA----- | 51 |
| Aplicación de modelado de temas con LDA ----- | 52 |
| Ejemplo práctico: Proyecto 16 (ver Anexo 16)----- | 52 |
| MÓDULO 4. Deep Learning y redes neuronales para NLP ----- | 53 |
| Introducción al Deep Learning (Aprendizaje profundo) ----- | 53 |
| Introducción a TensorFlow----- | 53 |
| Redes Neuronales Convolucionales CNN's----- | 54 |
| Redes Neuronales Recurrentes RNN's----- | 54 |
| Regresión Lineal con TensorFlow----- | 55 |
| Ejemplo práctico: Proyecto 17 (ver Anexo 17)----- | 55 |

| | |
|--|-----------|
| Clasificador de texto con TensorFlow | 55 |
| Modelado y TensorFlow ----- | 56 |
| Entropía Cruzada Binaria ----- | 56 |
| Estabilidad numérica----- | 56 |
| Aplicación de clasificador con TensorFlow | 57 |
| Ejemplo práctico: Proyecto 18 (ver Anexo 18) ----- | 57 |
| La Neurona | 57 |
| De la regresión a las Neuronas ----- | 58 |
| Neurona Artificial ----- | 58 |
| Aplicaciones de las neuronas artificiales ----- | 59 |
| ¿Cómo aprende un modelo? | 59 |
| Tasa de aprendizaje----- | 60 |
| Redes Neuronales Artificiales | 60 |
| Forward Propagation ----- | 61 |
| Importancia de las Redes Neuronales | 62 |
| Funciones de Activación | 62 |
| Función Sigmoidal ----- | 63 |
| Función Tangente ----- | 63 |
| Hiperbólica Función ReLu ----- | 63 |
| Función SoftMax ----- | 63 |
| Embeddings | 64 |
| Codificación One-Hot ----- | 64 |
| ¿Cómo funcionan los Embeddings en una red neuronal?----- | 65 |
| Geometría de los Embeddings ----- | 66 |
| Redes Neuronales Convolucionales CNN | 66 |
| Convolución de redes neuronales ----- | 67 |
| ¿Cómo funciona la convolución ----- | 67 |
| Pattern Matching (Emparejamiento de patrones)----- | 68 |
| Weight Sharing (Comparación de pesos)----- | 68 |
| Ventajas del Weight Sharing ----- | 69 |
| Convolución con imágenes a Color ----- | 69 |
| Filtros 3D----- | 69 |
| Arquitectura de una CNN----- | 70 |
| ¿Por qué es importante el pooling?----- | 70 |
| Convenciones y estrategias en la estructura CNN ----- | 71 |
| Pérdida de información ----- | 71 |
| Flexibilidad de la arquitectura ----- | 71 |
| CNN's para textos | 71 |
| Arquitectura en Textos ----- | 72 |
| Aplicaciones de las CNN's en NLP | 72 |

| | |
|---|-----------|
| Ejemplo práctico: Proyecto 21 (ver Anexo 21) ----- | 72 |
| Ejemplo práctico: Proyecto 22 (ver Anexo 22) ----- | 72 |
| Redes Neuronales Recurrentes RNN ----- | 73 |
| Ventajas de las RNN's ----- | 74 |
| Representación matemática de una RNN----- | 75 |
| Aplicaciones de las RNN's ----- | 75 |
| Aplicación de un clasificador de texto con RNN ----- | 76 |
| Ejemplo práctico: Proyecto 23 (ver Anexo 23) ----- | 76 |
| Ejemplo práctico: Proyecto 24 (ver Anexo 24) ----- | 76 |
| MÓDULO 5. Aplicaciones robustas con AI ----- | 76 |
| Ejemplo práctico: Proyecto 25 (ver Anexo 25) ----- | 76 |