

Inteligencia Artificial Aplicada para la Economía

Profesores

Profesor Magistral

Camilo Vega Barbosa

Asistente de Docencia

Sergio Julian Zona Moreno





Redes Neuronales Recurrentes para Texto


Del procesamiento secuencial a la
comprensión lingüística



Procesando las palabras y su lingüística

 Las palabras no están aisladas, sino que forman secuencias donde el orden importa. Al igual que no podemos entender una película viendo escenas aleatorias, una red neuronal necesita procesar el texto en orden para captar su significado. Las **RNN** son como lectores que recuerdan lo que leyeron anteriormente.

 Imagina a las RNN como personas leyendo un libro: cada palabra nueva se interpreta en el **contexto** de las anteriores. La frase "El banco está cerrado" tiene significado diferente si antes leíste sobre finanzas o sobre un parque. Las RNN mantienen esta **memoria contextual**.

 A diferencia de otros modelos que ven datos como fotos estáticas, las RNN tratan el texto como una **película en desarrollo**. Esta capacidad las hace ideales para traducción, análisis de sentimiento, o generación de texto, donde el **significado fluye** a través de toda la secuencia.



Texto como Secuencia: El Desafío

El procesamiento de texto presenta desafíos únicos que requieren modelos especializados como las RNN.



Características del Texto como Datos Secuenciales

- **Dependencia contextual:** El significado de una palabra depende de las que la rodean
- **Estructura jerárquica:** Caracteres → palabras → frases → párrafos → documentos
- **Longitud variable:** Desde mensajes cortos hasta documentos extensos
- **Gramática y sintaxis:** Reglas que rigen las relaciones entre palabras



¿Por qué las RNN para texto?

Las redes neuronales tradicionales no pueden capturar estas dependencias contextuales ni manejar secuencias de longitud variable.



RNN en Procesamiento de Lenguaje Natural



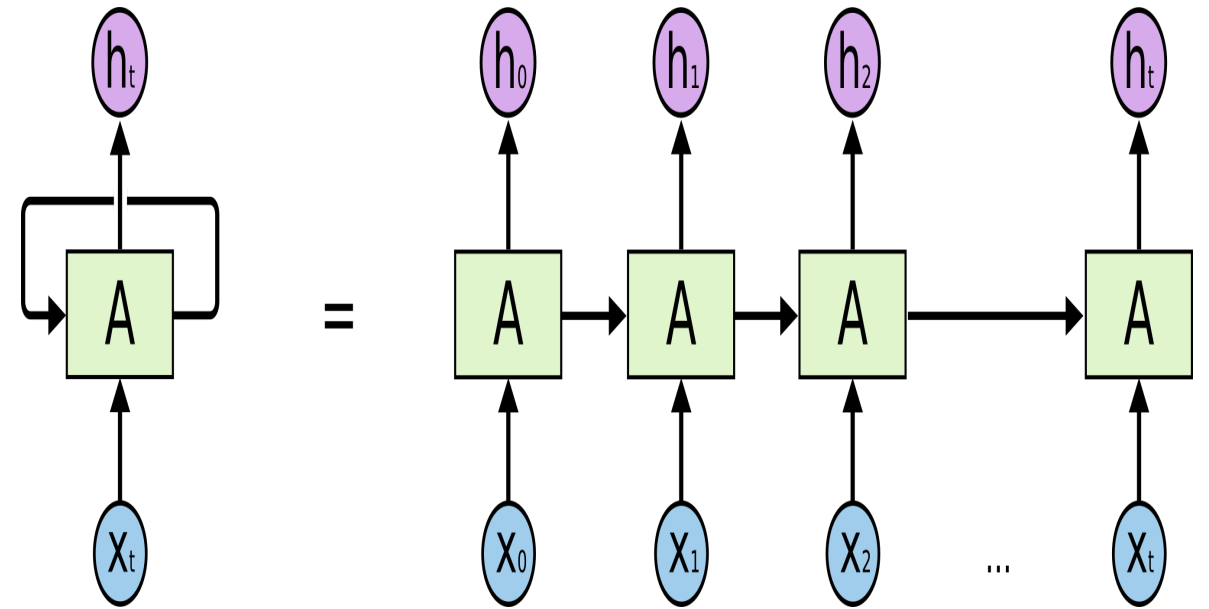
Aplicación en Texto

- **Unidades de entrada:** Palabras o tokens
- **Estado oculto:** Memoria del contexto previo
- **Aprendizaje:** Patrones lingüísticos y relaciones



Tareas Principales

- Clasificación de textos
- Generación de lenguaje
- Traducción automática
- Análisis de sentimiento



Tomado de: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.

A diferencia de las series temporales numéricas, con el texto necesitamos convertir palabras en vectores primero (embeddings).



Preparación de Datos Textuales para RNN

1 Tokenización

- División del texto en unidades (tokens)
- Puede ser por palabras, caracteres o subpalabras
- Construcción de vocabulario

2 Vectorización

- **Word embeddings:** Transformación de tokens a vectores
- **Embeddings pre-entrenados:** Word2Vec, GloVe, FastText
- **One-hot encoding:** Datos binarios

3 Secuenciación

- Organización en secuencias de longitud definida
- Padding para uniformizar longitudes
- División por oraciones o ventana deslizante

4 Etiquetado

- Asignación de objetivos según la tarea
- Etiquetas por secuencia (clasificación)
- Etiquetas por token (etiquetado gramatical)

Tipos de Etiquetado en RNN para Texto

El etiquetado es el proceso de asignar categorías o valores a los datos para que la RNN aprenda patrones.

1 Etiquetado a Nivel de Secuencia

- **Análisis de Sentimiento**
"Los inversores reaccionaron positivamente" → POSITIVO
- **Clasificación Temática**
"El banco anunció nuevas tasas" → FINANCIERO
- **Detección de Intención**
"¿Cuál será el PIB este trimestre?" → CONSULTA_PREDICCIÓN

2 Etiquetado a Nivel de Token

- **POS Tagging (Partes del Discurso)**
"El mercado creció rápidamente" → [DET, SUST, VERB, ADV]
- **NER (Entidades Nombradas)**
"Apple anunció una inversión" → [ORG, VERB, DET, SUST]
- **Segmentación**
"EstadosUnidos|reportó|crecimiento" → división de compuestos

Ejemplos Prácticos de Etiquetado Económico

El tipo de etiquetado determina qué aprenderá la RNN y qué tareas podrá realizar con textos económicos.

- **Análisis de Sentimiento Financiero:**

"Los resultados financieros **superaron las expectativas** de los analistas"

Etiqueta: [POSITIVO] → posible señal para decisiones de inversión

- **Reconocimiento de Entidades Económicas (NER):**

"El **BCE** anunció una tasa de interés del **0.5%** efectiva desde **enero 2023**"

Etiquetas: [INSTITUCIÓN, O, O, O, O, PORCENTAJE, O, O, FECHA]

- **Etiquetado de Relaciones:**

"**Amazon** adquirió **Whole Foods** por **13.7 mil millones**"

Relación: [Adquirente: Amazon, Adquirido: Whole Foods, Monto: 13.7B]



Ejemplo Dataset para RNN: Análisis de Sentimiento

ID	Texto Financiero	Etiqueta
1	"Beneficios de Tesla superaron expectativas"	POSITIVO
2	"Dólar se fortaleció tras anuncio de la Fed"	NEUTRAL
3	"Inflación aumentó, presionando mercados"	NEGATIVO
4	"Banco central mantuvo tasas sin cambios"	NEUTRAL
5	"Recortes de personal por pérdidas inesperadas"	NEGATIVO

Resultado del modelo entrenado: La RNN aprende a:

- **Clasificar** nuevos textos financieros en categorías de sentimiento
- **Asignar probabilidades** a cada posible clasificación (ej: 85% positivo)
- **Identificar patrones lingüísticos** asociados a movimientos de mercado



Ejemplo Intuitivo: RNN Procesando una Oración

Frase económica: "El banco central anunció un recorte inesperado de tasas de interés"

Análisis paso a paso:

1. La RNN lee "El banco" → Posibles contextos: institución financiera o asiento público
2. Añade "central" → Contexto financiero confirmado
3. Añade "anunció un recorte" → Detecta acción negativa potencial
4. Añade "inesperado" → Identifica sorpresa para mercados
5. Finaliza con "tasas de interés" → Clasifica como noticia financiera relevante

Cada palabra proporciona información adicional que actualiza la "memoria" de la red, permitiéndole refinar su comprensión del contexto económico.

Ejemplo Práctico: Análisis de Sentimiento

La RNN "lee" noticias financieras como un analista, acumulando contexto palabra a palabra.

Proceso de análisis:

1. **Contexto inicial:** Identificación del tema económico
"Los inversores reaccionaron positivamente al anuncio..."
2. **Memoria:** Acumulación de información relevante
"...sobre la fusión entre las principales empresas del sector..."
3. **Predicción:** Evaluación del impacto económico
"...lo que impulsó un aumento del 3% en el índice bursátil."

Resultado: Noticia con sentimiento positivo → Señal alcista para mercados

Limitaciones de las RNN en Texto

1 Pérdida de Memoria en Secuencias Largas

- **Problema:** La palabra 100 no "recuerda" la palabra 1
- **Ejemplo:** Al analizar un informe financiero extenso, la RNN olvida información crucial mencionada al inicio
- **Consecuencia:** Predicciones sesgadas hacia información reciente

2 Costos Computacionales Elevados

- Procesamiento secuencial (no paralelizable)
- Tiempos de entrenamiento prolongados para corpus extensos
- Dificultad para implementar en sistemas con recursos limitados

⚠ Más Limitaciones y Soluciones Parciales

3 El Problema del Contexto Unidireccional

- Las RNN tradicionales solo procesan texto en una dirección (izquierda a derecha)
- En "Las acciones subieron después de que la empresa anunciara pérdidas", el sentimiento depende de toda la frase
- **Solución parcial:** Arquitecturas bidireccionales (BiLSTM, BiGRU)

4 Dificultad con Ambigüedades

- **Problema:** "El banco quebró" (¿institución financiera o asiento físico?)
- **Consecuencia:** Interpretaciones incorrectas sin suficiente contexto
- **Solución parcial:** Atención y modelos más profundos

Arquitecturas RNN Especializadas para Texto

LSTM para Textos Largos

- Capta dependencias a larga distancia
- Ideal para documentos extensos, análisis de sentimiento
- Mantiene contexto entre oraciones

GRU para Eficiencia

- Balance entre rendimiento y velocidad
- Adecuado para clasificación y generación simple
- Recomendado cuando hay limitaciones de recursos

Bidireccional (Bi-LSTM/Bi-GRU)

- Procesa texto en ambas direcciones
- Captura contexto completo (anterior y posterior)
- Superior en tareas como etiquetado y extracción
- Base de modelos como BERT

Evolución: Hacia los Transformers

Transformers: La Nueva Generación

Los Transformers han superado a las RNN gracias a:

- **Atención paralela:** Procesamiento simultáneo de toda la secuencia
- **Atención bidireccional:** Consideran contexto completo en ambas direcciones
- **Escalabilidad:** Pueden manejar contextos más largos (aunque limitados)

Los **modelos basados en Transformers** como BERT y GPT dominan ahora el panorama del NLP, especialmente para tareas de alta complejidad lingüística y generación de texto.

Las RNN siguen siendo útiles para aplicaciones con secuencias cortas, recursos limitados o donde la interpretabilidad es crucial.

Recursos del Curso

Plataformas y Enlaces Principales

GitHub del curso

 github.com/CamiloVga/IA_Aplicada

Asistente IA para el curso

 [Google Notebook LLM](#)