**Cuestionario Evaluativo: Redes Neuronales e IA Aplicada**

**Juan Diego Gonzalez Paez**

**Inteligencia Artificial**

**Junio 2025**

**Pregunta 1 (10 puntos) - Fundamentos de CNN**

Explique qué son las redes neuronales convolucionales (CNN) y por qué son especialmente efectivas para el reconocimiento de imágenes. Mencione al menos tres características distintivas de las CNN que las hacen superiores a las redes neuronales tradicionales para tareas de visión por computadora.

R// las redes neuronales convolucionales son un tipo de red neuronal diseñadas para procesar datos con una estructura en forma cuadricula, como las imágenes. y son especialmente efectivas para el reconocimiento de imágenes ya que pueden aprender por ellas mismas características espaciales de las imágenes, lo que les ayuda a identificar patrones como texturas y formas complejas

**Tres características:**

**1.Invariancia espacial (traslación, escala y rotación):**

Gracias a la combinación de convolución y pooling, las CNN desarrollan cierta invariancia a pequeñas transformaciones espaciales, como traslaciones, escalas o rotaciones leves. Esto significa que pueden reconocer un objeto aunque esté en diferentes posiciones o tamaños dentro de la imagen.

**2.Submuestreo o pooling**:

Las capas de pooling reducen la dimensionalidad de la imagen manteniendo las características más relevantes. Esto hace a la red más robusta a pequeñas variaciones en la posición, orientación o escala de los objetos en la imagen.

**3.Mejor generalización con menos parámetros:**

A diferencia de una red neuronal totalmente conectada, una CNN no necesita conectar todos los píxeles con cada neurona.

**Pregunta 2 (10 puntos) - Arquitectura y Componentes**

Describa las capas principales que componen una CNN típica (Conv2D, MaxPooling2D, Flatten, Dense) y explique la función específica de cada una en el proceso de reconocimiento de imágenes. ¿Por qué es importante el orden de estas capas?

R//**1. Conv2D**

Su Función esta en Aplicar filtros a la imagen de entrada para detectar características locales, como bordes, texturas o formas. Cada filtro se desliza sobre la imagen y produce un mapa de activación que destaca dónde aparece ese patrón específico.

**2.MaxPooling2D**

Su Función es Reducir la dimensión espacial de los mapas de activación, manteniendo las características más importantes. Toma el valor máximo dentro de una ventana y descarta el resto, haciendo la red más eficiente y robusta frente a pequeñas variaciones.

**3. Flatten**

Su Función radica en Transforma la salida tridimensional de las capas en un vector unidimensional. Este paso es necesario para conectar la parte convolucional con la parte totalmente conectada. Prepara los datos para la toma de decisiones finales.

**4. Dense**

Su Función es que Cada neurona está conectada a todas las neuronas de la capa anterior. Estas capas combinan las características extraídas y aprenden relaciones no lineales complejas entre ellas.

El orden refleja una secuencia lógica de procesamiento de la imagen:

-Primero se extraen características locales (Conv2D).

-Luego se resumen y simplifican esas características (MaxPooling2D).

-Después se preparan los datos para la toma de decisiones (Flatten).

-Finalmente, se clasifica o predice el resultado basado en lo aprendido (Dense).

**Pregunta 3 (12 puntos) - Preprocesamiento de Datos**

En el contexto del dataset CIFAR-10: a) ¿Por qué es necesario normalizar los valores de píxeles al rango [0, 1]? (4 puntos) b) ¿Qué significa convertir las etiquetas a formato "one-hot" y por qué es necesario? (4 puntos) c) Mencione dos técnicas de data augmentation que podrían mejorar el rendimiento del modelo y explique cómo funcionan. (4 puntos)

R// **a)** Normalizar los valores de los píxeles al rango [0,1] permite que el modelo entrene **más rápido y de forma más estable**. Esto se debe a que las redes neuronales funcionan mejor cuando los datos de entrada están en una **escala similar**.

**b)** Convertir las etiquetas al formato **"one-hot"** significa transformar cada etiqueta de clase en un vector binario en el que solo una posición tiene el valor 1 y las demás son 0. es necesario ya que muchos modelos de clasificación, comparan su salida con vectores one-hot durante el cálculo de la **función de pérdida categórica**. Sin este formato, el modelo no podría aprender a distinguir entre clases correctamente.

**c)** dos técnicas de data augmentation que podrían mejorar el rendimiento del modelo podrían ser:

1.**Desplazamiento horizontal:**  
Se mueve la imagen hacia la izquierda o derecha unos pocos píxeles.

**2.Giro aleatorio:**   
Esta técnica consiste en rotar las imágenes en ángulos pequeños, Esto ayuda al modelo a aprender que los objetos siguen siendo los mismos, aunque estén ligeramente girados.

**Pregunta 4 (10 puntos) - Optimización y Entrenamiento**

Analice los siguientes aspectos del entrenamiento de una CNN:

**¿Qué función de pérdida se utiliza para clasificación multiclase y por qué?**

R// se utiliza la función de pérdida llamada "categorical crossentropy". Esta función mide la diferencia entre la distribución de probabilidad real y la distribución predicha por la red

**¿Cuál es la diferencia entre usar 'adam' y 'sgd' como optimizadores?**

R// ADAM Es un optimizador que ajusta automáticamente la tasa de aprendizaje para cada parámetro y sgd Es más simple y requiere definir una tasa de aprendizaje fija, en conclusión, Adam es más rápido y automático por otro lado sgd, aunque sea más manual puede ofrecer mejores resultados y más estables si se configura bien

**¿Cómo se puede detectar y prevenir el overfitting durante el entrenamiento?**

se detecta cuando el modelo empieza a mejorar en los datos de entrenamiento, pero empeora en los datos de validación. Esto indica que está memorizando en lugar de generalizar. Para prevenirlo se pueden usar varias estrategias, entre ellas esta:

-Regularización: apagar aleatoriamente algunas neuronas durante el entrenamiento, obligando a la red a no depender de caminos específicos.

-Early stopping: detener el entrenamiento cuando la pérdida de validación deja de mejorar.

**Pregunta 5 (10 puntos) - Transfer Learning**

El taller menciona el uso de MobileNetV2 pre-entrenado. Explique:

**¿Qué es transfer learning y cuáles son sus ventajas?**

R// Transfer learning consiste en usar un modelo ya entrenado (como MobileNetV2) en otro problema similar. La ventaja es que ahorra tiempo y recursos, porque el modelo ya aprendió a detectar patrones generales que sirven también para nuevas tareas

**¿Por qué se eligió MobileNetV2 para este proyecto específico?**

R// Se eligió **MobileNetV2** porque es un modelo **ligero, rápido y eficiente**, ideal para usar en computadoras comunes o dispositivos móviles.

**¿En qué situaciones sería preferible entrenar un modelo desde cero versus usar transfer learning?**

R//se recomienda entrenar un modelo desde cero cuando, Se tiene un dataset muy grande y específico o El problema es muy diferente al que resolvía el modelo preentrenado.

**Pregunta 6 (12 puntos) - Procesamiento de Lenguaje Natural**

Respecto al componente NLP del sistema integrado: a) Explique qué es la lemmatización y por qué es importante en el procesamiento de texto. (4 puntos) b) ¿Cómo funcionan los patrones de conversación definidos en el código para identificar intenciones del usuario? (4 puntos) c) Mencione tres técnicas que podrían mejorar la capacidad de comprensión del chatbot. (4 puntos)

R//**a)** La lemmatizacion es el proceso de reducir una palabra a su forma base, que puede ser sumamente útil para la agrupación en clúster. Su **importancia en el procesamiento de texto** radica en varios puntos clave los cuales son mejorar la precisión, la optimización de recursos y la normalización de datos

**b)** Los patrones de conversación definidos en el código, funcionan como **filtros** que intentan asociar frases o palabras clave del usuario con una intención predefinida. Su funcionamiento se basa en la correspondencia entre la entrada del usuario y estos patrones.

**c)** **Aprendizaje Activo:** es una estrategia donde el chatbot, cuando no está seguro de la intención del usuario o de cómo clasificar una entrada, pide aclaraciones al usuario o aprende de las interacciones humanas.

**Manejo avanzado de contexto:** guardar y utilizar información de turnos anteriores para interpretar mejor las preguntas y dar respuestas coherentes.

**Incorporación de modelos de lenguaje contextualizados:** como BERT o GPT, que entienden mejor el significado según el contexto de la conversación

**Pregunta 7 (10 puntos) - Integración de Sistemas**

El taller propone integrar reconocimiento de imágenes con NLP. Describa:

**¿Cuáles son los principales desafíos técnicos de esta integración?**

R// **Gestión de ambigüedades:** A veces, la imagen puede ser ambigua o no contener información clara, lo que dificulta que el chatbot responda adecuadamente sin un contexto textual claro.

**Disponibilidad de Datos Anotados:** La escasez y el alto costo de crear grandes conjuntos de datos de imágenes y descripciones textuales alineadas.

**Interpretación coherente de los datos:** Los datos visuales y textuales son muy diferentes, por lo que combinarlos para que el sistema entienda ambos contextos simultáneamente es complejo.

**¿Cómo se mantiene el contexto entre el análisis de imágenes y la conversación?**

R// La información del contexto se conserva a través de un sistema de administración de diálogos que almacena y vincula datos tanto del reconocimiento de imágenes como de las interacciones anteriores del usuario.

**Proponga una mejora específica para hacer más fluida esta integración.**

R// Implementar un sistema de memoria contextual persistente que almacene no solo la información inmediata de la imagen y el texto, sino también el historial completo de la conversación y análisis previos. Esto permite que el chatbot recuerde detalles importantes a lo largo del tiempo y pueda hacer referencias cruzadas entre imágenes y diálogos pasados, ofreciendo respuestas más coherentes y personalizadas.

**Pregunta 8 (8 puntos) - Análisis de Rendimiento**

Para evaluar el desempeño de una CNN: a) ¿Qué información proporciona una matriz de confusión? (4 puntos) b) ¿Cuál es la diferencia entre accuracy, precision y recall? ¿Cuándo es más importante cada métrica? (4 puntos)

R// a) La matriz de confusión muestra cómo se desempeña un modelo de clasificación, al comparar las predicciones del modelo con las etiquetas reales, la matriz de confusion nos proporciona la información de Dónde se cometieron errores, Cuántos ejemplos fueron clasificados correctamente y entre qué clases se confunde el modelo.

b) **Accuracy:** Proporción total de predicciones correctas. Útil cuando las clases están balanceadas y todos los errores tienen el mismo costo.

**Precision:** De las veces que el modelo predijo una clase, cuántas veces acertó. Importante cuando los falsos positivos son costosos

**Recall:** De todas las veces que una clase debería haber sido detectada, cuántas veces lo fue. Clave cuando los falsos negativos son críticos (por ejemplo, no detectar un objeto peligroso).

**Pregunta 9 (10 puntos) - Casos de Uso Específicos**

El sistema se plantea para apoyo académico y consultas psicológicas básicas. Analice:

**¿Qué consideraciones éticas y de privacidad se deben tener en cuenta?**

R// Límites de Competencia: El sistema debe operar dentro de sus capacidades predefinidas y no pretender reemplazar la intervención profesional

**Consentimiento informado:** El usuario debe saber qué datos se recopilan y cómo se usan.

Privacidad de datos: Toda información que el usuario comparta (imágenes, texto) debe manejarse bajo estrictas políticas de protección, cumpliendo normativas como la Ley de Protección de Datos Personales.

**¿Cómo se podría adaptar el sistema para detectar situaciones que requieran intervención humana?**

R//El sistema puede adaptar un módulo de lenguaje emocional con la función de detectar frases asociadas con la depresión o ansiedad. activándose este patrón el sistema puede generar y hacer sonar una alarma para el llamado de una persona profesional

**Proponga una funcionalidad adicional que agregue valor al sistema.**

R//una funcionalidad adicional podría ser el diario emocional inteligente, que trata de que el usuario va agregando su día a día de cómo le fue, como se sintió, que actividades realizo y asi el sistema analice estos registros y de tal manera pueda detectar patrones emocionales y Sugerir actividades académicas o de bienestar personal según el estado emocional detectado.

**Pregunta 10 (8 puntos) - Visión Futura**

Considerando las extensiones propuestas en el taller:

**¿Cuál de las extensiones mencionadas considera más importante implementar primero y por qué?**

R//yo considero más importante la integración de un modelo de lenguaje más sofisticado (como BERT o GPT) ya que Mejora la comprensión del lenguaje natural, lo que permite interpretar mejor tanto preguntas académicas como expresiones emocionales sutiles.

**¿Cómo impactarían los avances en modelos de lenguaje como GPT o BERT en este tipo de sistemas?**

R// Modelos como GPT o BERT pueden mejorar significativamente estos sistemas ya que Entenderían mejor el lenguaje natural, incluyendo frases ambiguas, emocionales o informales. Permitiendo mantener conversaciones más coherentes y contextuales, recordando lo que el usuario dijo antes.