1. Introducción

El presente informe analiza modelos de clasificación (1D/tabular, 2D/imágenes, late-fusion y early-fusión) desarrollados para abordar un problema de clasificación, enfrentando un desafío común de sobreajuste. Se implementaron diversas estrategias de regularización para mitigar este fenómeno y mejorar la generalización de los modelos.

2. Modelo Tabular: Análisis de Hiperparámetros y Regularización

2.1 Descripción del Modelo

* Red neuronal con capas densas
* Objetivo: Clasificación de datos tabulares

2.2 Técnicas de Regularización

1. **Regularización L2**
   * Penalización de pesos grandes
   * Prevención de complejidad excesiva
2. **Early Stopping**
   * Detención del entrenamiento cuando no hay mejora
   * Restauración de mejores pesos

2.3 Resultados de Hiperparámetros

| **Tasa de Aprendizaje** | **Tamaño del Lote** | **Precisión Entrenamiento** | **Precisión Validación** | **Precisión Prueba** |
| --- | --- | --- | --- | --- |
| 0.001 | 32 | 0.7471 | 0.6470 | 0.6906 |
| 0.005 | 32 | 0.7061 | 0.6262 | 0.6683 |
| 0.01 | 32 | 0.6997 | 0.6054 | 0.6603 |

2.4 Análisis de Resultados

* Tasa de aprendizaje más baja (0.001) mostró mejor rendimiento
* Tamaño de lote pequeño (32) más efectivo
* Regularizaciones ayudaron a controlar el sobreajuste

3. Modelo de Imágenes: Arquitectura ResNet50 y Regularización

3.1 Modelos Desarrollados

1. **Modelo 1: Transfer Learning**
   * Congelamiento total de capas
   * Precisión entrenamiento: 80.83%
   * Precisión validación: 55.43%
2. **Modelo 2: Transfer Learning con Batch Size Diferente**
   * Batch size: 128
   * Precisión entrenamiento: 88.50%
   * Precisión validación: 52.56%
3. **Modelo 3: Fine Tuning**
   * Descongelamiento de últimas 10 capas
   * Precisión entrenamiento: 93.72%
   * Precisión validación: 48.72%

3.2 Técnicas de Regularización

1. **Regularización L2**
2. **Dropout**
3. **Transfer Learning**
   * Congelamiento/descongelamiento selectivo de capas

3.3 Análisis Comparativo

| **Modelo** | **Precisión Entrenamiento** | **Precisión Validación** | **Precisión Prueba** | **Nivel de Sobreajuste** |
| --- | --- | --- | --- | --- |
| Modelo 1 | 80.83% | 55.43% | 51.99% | Moderado |
| Modelo 2 | 88.50% | 52.56% | 52.95% | Alto |
| Modelo 3 | 93.72% | 48.72% | 46.73% | Muy Alto |

4. Modelos de Fusión: Mitigación del Sobreajuste

4.1 Modelos Individuales

* **Modelo Tabular**
  + Precisión prueba: 0.6906
* **Modelo de Imágenes**
  + Precisión prueba: 0.5295

4.2 Estrategias de Fusión y Regularización

1. **Fusión Tardía (Late Fusion)**
   * Grid Search con Random Forest
   * Validación Cruzada
   * Weighted Average Ensemble
2. **Fusión Temprana (Early Fusion)**
   * Capas de Dropout
   * Early Stopping

4.3 Resultados de Fusión

| **Método de Fusión** | **Precisión Prueba** |
| --- | --- |
| Modelo Tabular Individual | 0.6906 |
| Weighted Average Ensemble | 0.6683 |
| Fusión Temprana | 0.6411 |
| Grid Search con Random Forest | 0.5901 |
| Validación Cruzada con Random Forest | 0.5885 |
| Modelo de Imágenes Individual | 0.5295 |

4.4 Análisis comparativo

**Rendimiento general**:

* + Modelo Tabular individual: 0.6906 (mejor rendimiento en prueba)
  + Weighted Average Ensemble (Late Fusion): 0.6683 (Validation-Modelos Base )
  + Early Fusion: 0.6411 (Train-Modelos Base)
  + Otros enfoques de Late Fusion: < 0.60 (Train-Modelos Base)

**Comparación de fusión temprana vs. tardía**:

* + La fusión temprana (0.6411) supera a la mayoría de los enfoques de fusión tardía, sin embargo, no logra superar al modelo tabular individual ni al Weighted Average Ensemble.

**Influencia del conjunto de datos de entrenamiento**:

* + Early Fusion: Utiliza características extraídas de los modelos base en el conjunto de entrenamiento, lo que puede llevar a incorporar el sobreajuste del modelo de imágenes.
  + Late Fusion (Grid Search y Validación Cruzada): También se ven afectados por el uso de predicciones de entrenamiento.
  + Weighted Average Ensemble: Se beneficia del uso de predicciones de validación para calcular los pesos.

**Complejidad vs. Rendimiento**:

* + La fusión temprana introduce una complejidad adicional (red neuronal con capas densas) pero no logra superar los métodos más simples como el modelo tabular individual o el Weighted Average Ensemble.

**Mitigación del sobreajuste**:

* + Early Fusion: Intenta mitigar el sobreajuste con capas de Dropout y Early Stopping, pero aun así no supera al modelo tabular individual.
  + Weighted Average Ensemble: Logra una mejor mitigación del sobreajuste al basarse en el rendimiento de validación.

5. Razones de las diferencias en rendimiento

1. **Propagación del sobreajuste**:
   * Tanto en Early Fusion como en algunos métodos de Late Fusion, el sobreajuste del modelo de imágenes en el conjunto de entrenamiento se propaga al modelo final.
2. **Capacidad de generalización**:
   * El modelo tabular individual y el Weighted Average Ensemble muestran una mejor capacidad de generalización..
3. **Balanceo de influencias**:
   * El Weighted Average Ensemble logra un mejor balance entre los modelos base, mientras que la Early Fusion puede estar dando demasiado peso a características menos informativas.

6. Conclusiones y recomendaciones finales

1. **Uso inteligente de los datos**:
   * El modelo tabular individual y el Weighted Average Ensemble muestran los mejores resultados, sugiriendo que el uso inteligente de los datos de validación son clave.
2. **Reevaluación de la fusión temprana**:
   * Considerar técnicas para reducir la influencia del modelo de imágenes sobreajustado en la fusión temprana, como la regularización de características o la selección de características más robustas.
3. **Mejora del modelo de imágenes**:
   * Sigue siendo una prioridad mejorar este modelo para reducir el sobreajuste y potencialmente aumentar su contribución positiva en cualquier esquema de fusión.
4. **Exploración de técnicas híbridas**:
   * Investigar métodos que combinen aspectos de fusión temprana y tardía, posiblemente con una selección dinámica de características o modelos basada en el rendimiento de validación.
5. **Análisis detallado de errores**:
   * Examinar los casos donde cada enfoque de fusión falla o acierta para entender mejor sus fortalezas y debilidades específicas.

En resumen, aunque la fusión temprana muestra un rendimiento razonable, no logra superar al modelo tabular individual ni al Weighted Average Ensemble. Esto sugiere que, para este problema específico, los enfoques que hacen un uso inteligente de los datos de validación son más efectivos. La clave para mejorar el rendimiento general parece estar en refinar el modelo de imágenes y en desarrollar métodos de fusión que puedan aprovechar las fortalezas de ambos modelos base de manera más efectiva.