

OPTIMIZACIÓN DE DECISIONES DE 4^{ta} OPORTUNIDAD EN FÚTBOL AMERICANO PROFESIONAL: UN ENFOQUE DE APRENDIZAJE MULTI-TAREA

Marco Villagrán, Autor. Oscar Acosta, Profesor.
Diplomado en Ciencia de Datos FES Acatlán, Gen. 27,
Universidad Nacional Autónoma de México, Ciudad de México, México

Abstract

Las decisiones en 4^a oportunidad influyen de forma desproporcionada en la probabilidad de victoria (WP) de un equipo. En este trabajo se presenta un modelo que estima el cambio de WP por acción (ΔWP) para *go-for-it* (jugársela), *field goal* (intento de gol de campo) y *punt* (despeje). Se utilizaron jugadas de 4^a oportunidad de las temporadas 2016–2024 de la NFL ($n \approx 37,300$) etiquetadas con $nfl14th$ (*teacher* de w_{pgo} , w_{pfg} , w_{ppunt}) y se construyó un MLP multi-tarea con cabezas auxiliares para p (FG make) y p (4th conv). El conjunto de variables combina estado del juego e ingeniería situacional (distancia de FG, indicadores de *late & close*, diferenciales de tiempos fuera). Se entrenó con partición por temporadas (train: 2016–2021; val: 2022; test: 2023–2024) y pérdida Huber sobre ΔWP . En test se alcanzó **0.794** de *policy agreement* y **0.23 %** de *mean regret*; en escenarios *late & close* se consiguió **0.744** y **0.99 %**, respectivamente. El análisis por bins (yarda y *yards-to-go*) muestra mayor utilidad en medio campo y 3–6 yardas, donde la decisión es menos clara. Los resultados sugieren que un MLP bien regularizado y con *features* específicas captura con fidelidad las compensaciones clave de 4^a oportunidad, ofreciendo una base reproducible para evaluación y comparación futura.

INTRODUCCIÓN

Las decisiones en 4^a oportunidad concentran un impacto desproporcionado en la probabilidad de victoria (WP) de un equipo. Elegir entre *jugársela* (go-for-it), intentar un *field goal* (FG) o realizar un *punt* (despeje) altera el estado del partido de forma inmediata y condiciona secuencias posteriores. Aunque existen reglas empíricas y heurísticas ampliamente difundidas, tomar esta decisión requiere de un gran contexto situacional: tiempo restante, distancia por avanzar, posición de campo, diferencia en el marcador, tiempos fuera disponibles y condiciones ambientales, entre otros factores. Esta variabilidad abre la puerta a sistemas de apoyo a la decisión basados en datos que estimen de manera consistente el efecto marginal de cada acción.

En este trabajo se estima *el ΔWP (win percentage o diferencial de Probabilidad de Victoria) por acción*. En lugar de modelar WP absoluta y derivar recomendaciones por diferencia, se estiman directamente tres salidas continuas: ΔWP_{go} , ΔWP_{fg} y ΔWP_{punt} , definidas como la diferencia entre la WP posterior a ejecutar cada acción y la WP del estado actual. Para construir etiquetas de entrenamiento

se emplearon probabilidades de victoria por acción proporcionadas por $nfl14th$ (*teacher*), a partir de las cuales se derivaron las variaciones ΔWP . De esta forma se alinea la función objetivo con la métrica de decisión.

Se trabajó con todas las jugadas de 4^a oportunidad de las temporadas 2016–2024 de la NFL ($n \approx 37,300$). El conjunto de variables incluye estado del juego (cuarto, tiempo restante por partido y por mitad, yarda del campo, distancia por avanzar, diferencia en el marcador, tiempos fuera), contexto (tipo de estadio y superficie) y factores ambientales (temperatura, viento, línea de apuesta). Se aplicaron normalizaciones e imputaciones conservadoras (p. ej., condiciones constantes para estadios cerrados) y se documentó el tratamiento de faltantes estructurales en w_{ppunt} , faltantes únicamente cuando es claramente la peor de las opciones. Para evaluar generalización temporal, se utilizó una partición por temporadas (entrenamiento: 2016–2021; validación: 2022; prueba: 2023–2024).

Metodológicamente, se entrenó un perceptrón multicapa (MLP) multi-tarea con una cabeza principal para las tres salidas ΔWP y cabezas auxiliares que predicen p (FG make) y p (4th conv) (probabilidad de éxito para *intento de gol de campo* y *jugada en 4^{ta} oportunidad*). Estas tareas auxiliares actúan como regularizadores que favorecen establecer límites posibles para la conversión de una cuarta oportunidad. Se empleó pérdida Huber sobre ΔWP , normalización de variables numéricas y codificación categórica de *roof/surface*. La evaluación se centra en dos métricas principales: *policy agreement* (coincidencia entre la acción con mayor ΔWP estimada y la acción con mayor ΔWP del *teacher*) y *mean regret* (diferencia entre la mejor ΔWP y la ΔWP de la acción elegida por el modelo), tanto de forma global como por bins de posición de campo y distancia por avanzar, con énfasis en escenarios *late & close*.

Las contribuciones principales son:

- Un pipeline reproducible de extremo a extremo (extracción, etiquetado, limpieza e ingeniería de variables) para 4^a oportunidad en NFL (2016–2024).
- Un modelo MLP multi-tarea que estima directamente las tres ΔWP por acción, con tareas auxiliares calibradas al problema.
- Un método de evaluación orientado a la decisión: métricas globales y por bins, además de un análisis específico de *late & close*.

- Resultados que muestran alto acuerdo de decisiones y bajo *regret* (arrepentimiento) promedio, junto con identificación de zonas de mayor utilidad (medio campo y 3–6 yardas).

DATOS

Fuentes y cobertura

Se emplearon datos públicos de `nflfastR` (jugada a jugada) y el paquete `nfl4th` para el etiquetado de 4ª oportunidad. La muestra incluye todas las jugadas de 4ª oportunidad (temporada regular y posttemporada) de la NFL entre 2016 y 2024, con un total de $N = 37,334$ ejemplos tras limpieza. La distribución por temporada es estable a lo largo del periodo.

Extracción y preprocesamiento

A partir de los archivos `pbp` se construyó un *pipeline* que filtró únicamente jugadas de 4ª oportunidad en temporada regular o posttemporada, extrajo variables de estado como cuarto, tiempos restantes, `yards_left` (posición en el campo), `ydstogo`, diferencial de marcador y tiempos fuera, la línea *spread* y se generó variables derivadas como recepción del *kickoff* de segunda mitad y distancia de gol de campo. Se normalizaron categorías de techo y superficie, se imputaron condiciones climáticas según el tipo de estadio, y finalmente se convirtieron variables a formato numérico.

Etiquetado (teacher) y objetivos

Se aplicó `nfl4th::add_4th_probs` para obtener WP condicional por acción: wp_{go} , wp_{fg} , wp_{punt} . A partir de ello se definieron los objetivos continuos

$$\Delta WP_{act} = wp_{act} - wp_{actual}, \quad act \in \{go, fg, punt\}.$$

Se observó faltante estructural en wp_{punt} ; en lugar de descartar ejemplos, se imputó un *piso constante* (mínimo global observado) y se documentó este caso como faltante en casos donde claramente era la peor opción.

Conjunto final y partición

Tras limpieza no quedaron valores nulos en las variables de modelado. La partición temporal fue: entrenamiento (2016–2021), validación (2022) y prueba (2023–2024), evitando fuga de información entre temporadas.

Variables empleadas

La siguiente lista resume las variables utilizadas. Además, se incorporaron indicadores situacionales (calculados) usados en el modelo (detallados en Metodología).

- `quarter`: Cuarto del partido (1–5/6 en OT).
- `game_seconds_remaining`: Segundos restantes en el partido.
- `half_seconds_remaining`: Segundos restantes en la mitad.

- `yardline_100`: Yarda relativa (0=TD rival, 100=propia yarda 0).
- `ydstogo`: Yardas por avanzar en 4ª oportunidad.
- `score_differential`: Marcador (ataque – defensa).
- `timeouts_off`, `timeouts_def`: Tiempos fuera restantes (ambos equipos).
- `receive_2h_ko`: Indicador de si recibe el *kickoff* de la 2ª mitad.
- `roof`: Tipo de estadio (*open/closed*).
- `surface`: Superficie (*grass/turf*).
- `temp_f`, `wind_mph`: Temperatura y viento (imputados).
- `spread_line`: Línea de apuesta previa al partido.
- `fg_dist_yd`: Aproximación de distancia de gol de campo (`yardline + 17`).
- ΔWP_{go} , ΔWP_{fg} , ΔWP_{punt} : Objetivos continuos por acción.

Calidad de datos

Antes de imputación, `temp_f` y `wind_mph` presentaban ~3.4% de faltantes (no calculados para en estadios cerrados) y `surface` ~1.7%. El faltante en wp_{punt} se concentró en posiciones de campo < 30 yardas del rival. Tras las imputaciones y normalizaciones descritas, el conjunto quedó libre de nulos y listo para entrenamiento y evaluación.

ANÁLISIS EXPLORATORIO DE DATOS (EDA)

El EDA se enfocó en (i) validar que las etiquetas del *teacher* (`nfl4th`) y las variables derivadas sean coherentes con el deporte y (ii) caracterizar las zonas del espacio de estados donde la decisión es menos clara. Se emplearon agregaciones por *bins* y comprobaciones de monotonía: la probabilidad de anotar un FG debe decrecer con la distancia de intento, y la probabilidad de convertir 4ª debe decrecer con las *yards-to-go*. Asimismo, se cuantificó la tasa de “decisiones similares” (diferencia entre la mejor y la segunda mejor $\Delta WP < 0.5\%$), útil para identificar escenarios de alta incertidumbre.

Probabilidades auxiliares: coherencia monótona

La Figura 1 muestra una relación decreciente y suave entre la distancia de FG y la probabilidad de acierto, con punto de inflexión alrededor de 45–50 yardas y caída pronunciada a partir de 55–60 yardas. De forma análoga, la Figura 2 confirma la relación decreciente entre *yards-to-go* y la probabilidad de conversión en 4ª, con valores altos (~0.7–0.9) en 4ª y 1–2 y descensos progresivos hacia 10+ yardas. Estas comprobaciones apoyan la validez de las etiquetas auxiliares y del preprocesamiento.

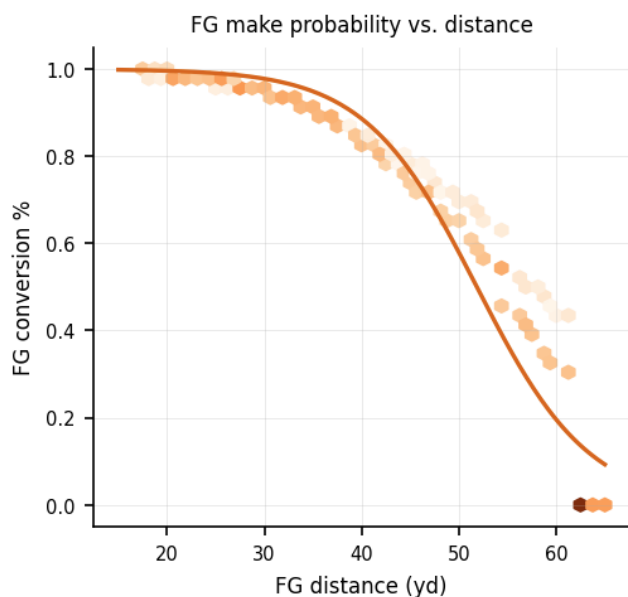


Figure 1: Probabilidad de acierto de FG vs. distancia de intento. Relación decreciente esperada y transición marcada en 45–50 yd.

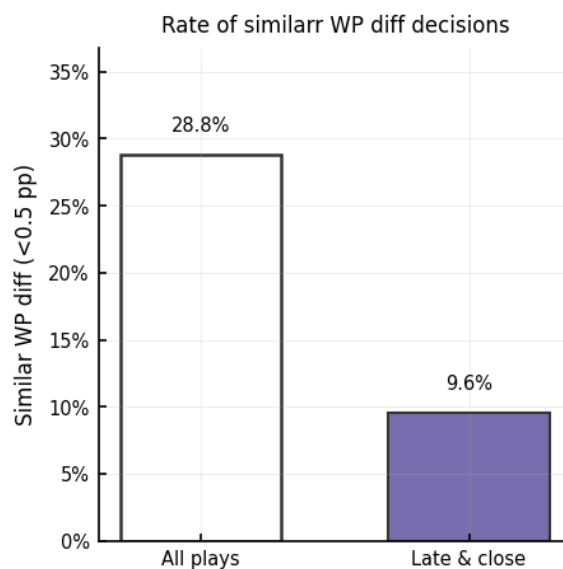


Figure 3: Tasa de decisiones con diferencia de $\Delta WP < 0.5$ % (diferenciales similares) en todo el conjunto y en *late & close*.

Patrones por posición de campo y yards-to-go

La Figura 4 presenta la *mejor acción* por bins de posición de campo: en red zone (0–20) el *teacher* favorece FG y go; entre 21–40 FG mantiene ventaja; en 41–60 la preferencia migra hacia go; y en campo propio profundo (61–100) domina *punt*.

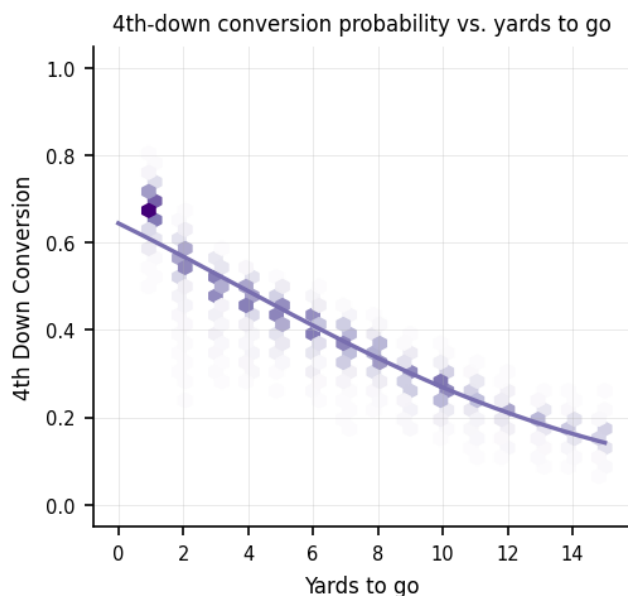


Figure 2: Probabilidad de conversión de 4ª vs. yards-to-go. Descenso monótono con la distancia por avanzar.

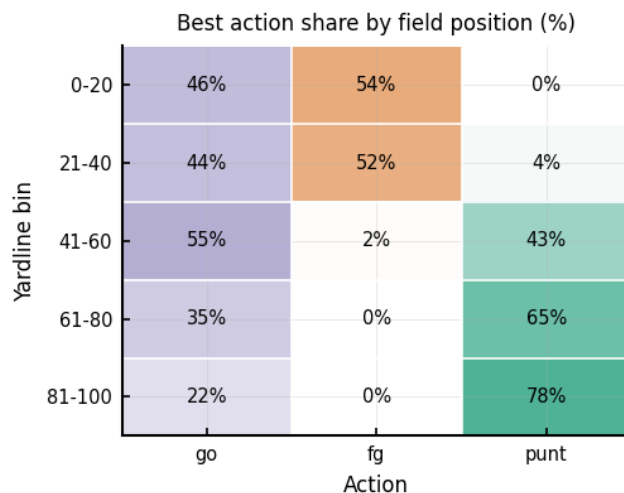


Figure 4: Proporción de mejor acción por bins de posición de campo (0–20, 21–40, 41–60, 61–80, 81–100).

Competitividad de la decisión

La Figura 3 resume la *tasa de decisiones similares* (margen entre la mejor y la segunda mejor $\Delta WP < 0.5$ %). En el conjunto completo se observó ~28.8%, mientras que en escenarios *late & close* (restan < 600 s y diferencia ≤ 8) la tasa se reduce a ~9.6%. Es decir, cuando el partido está apretado y cerca del final, el *teacher* tiende a concentrar más claramente la recomendación en una acción.

La Figura 5 muestra el patrón por yards-to-go: *go* prevalece en 1–2 yardas; *punt* aumenta con distancias largas; y FG gana peso cuando la distancia de intento es razonable aún con 7–10 yardas por avanzar. Estos mapas de calor sitúan las regiones donde el modelo es potencialmente más útil (medio campo y 3–6 yds), coherentes con los resultados posteriores.

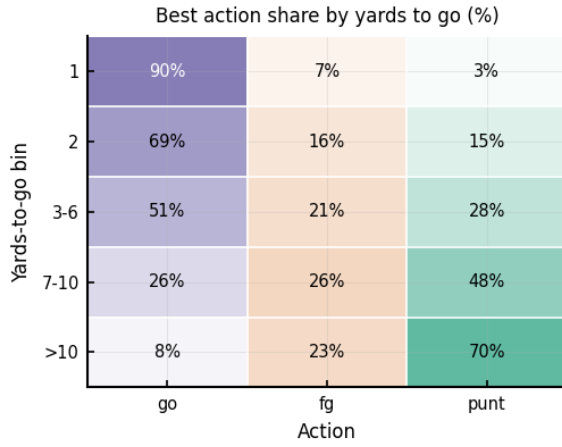


Figure 5: Proporción de mejor acción por *bins* de *yards-to-go* (1, 2, 3–6, 7–10, >10).

La Figura 6 muestra la frecuencia con que la decisión del *coach* difiere de la acción recomendada por el *teacher* en situaciones de 4ª oportunidad. En general, el desacuerdo es notable incluso en escenarios *late & close*, donde cabría esperar mayor alineación con estrategias óptimas, lo que sugiere un margen importante para mejorar la toma de decisiones desde un enfoque analítico.

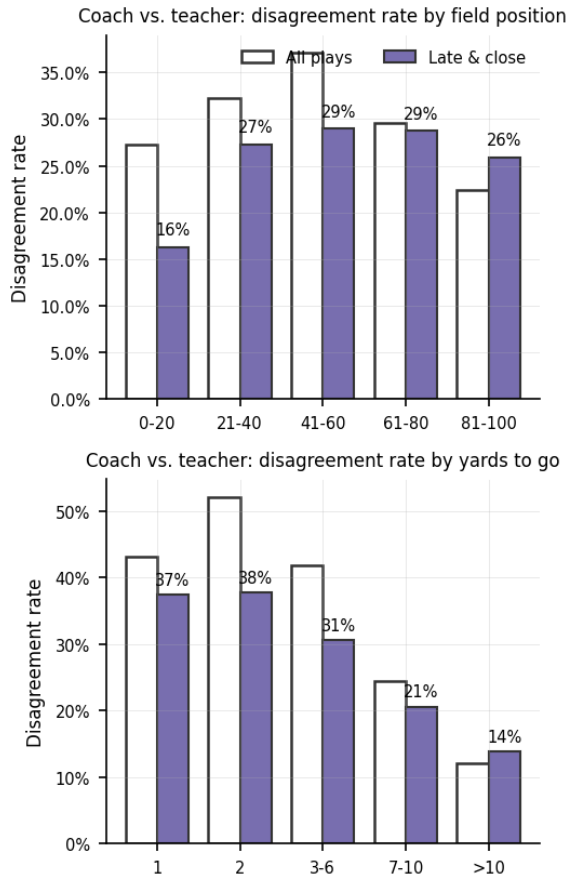


Figure 6: Tasa de desacuerdo entre *coach* y *teacher* en la acción seleccionada para 4ª oportunidad, considerando todas las jugadas y el escenario *late & close*.

Resumen. El EDA confirma que (i) las probabilidades auxiliares cumplen relaciones monótonas esperadas, (ii) la competitividad de la decisión varía con el contexto (menor ambigüedad en *late & close* ya que las decisiones son más determinantes) y (iii) los patrones por campo y distancia son consistentes con la intuición del deporte. Estas evidencias fundamentan a las etiquetas y al espacio de *features* para entrenar el modelo de ΔWP .

METODOLOGÍA

Formulación del problema

Sea x el estado del juego en 4ª oportunidad y $\mathcal{A} = \{\text{go}, \text{fg}, \text{punt}\}$ el conjunto de acciones. A partir del *teacher* (`nf14th`) se obtienen $w p_a(x)$ y $w p_{\text{actual}}(x)$, y se definen las etiquetas continuas

$$\Delta W P_a(x) = w p_a(x) - w p_{\text{actual}}(x), \quad a \in \mathcal{A}. \quad (1)$$

El modelo predice el vector $\hat{\mathbf{y}}(x) = [\hat{\Delta W P}_{\text{go}}, \hat{\Delta W P}_{\text{fg}}, \hat{\Delta W P}_{\text{punt}}]$ y la recomendación se induce por

$$\pi_{\text{modelo}}(x) = \arg \max_{a \in \mathcal{A}} \hat{\Delta W P}_a(x). \quad (2)$$

Arquitectura

La Figura 7 resume el flujo: (i) **entradas** numéricas y categóricas; (ii) un bloque de **preprocesamiento en red** (concatenación, normalización para numéricas y codificación *one-hot* para categóricas); (iii) un **bloque** denso apilado; y (iv) **tres cabezas** de salida (una principal para ΔWP y dos auxiliares).

- **Entradas y preprocesamiento.** Las variables numéricas se normalizan con estadísticas del conjunto de entrenamiento; *roof* y *surface* se codifican *one-hot*. Las binarias (p. ej., *receive_2h_ko*) se tratan como numéricas en $[0, 1]$.
- **Bloque denso.** Cuatro capas densas de $256 \rightarrow 128 \rightarrow 64 \rightarrow 32$ unidades con activación GELU (en nuestros experimentos, ReLU y GELU rindieron de forma similar), seguidas de BatchNorm, Dropout (0.10–0.15) y regularización L2 (10^{-5}).
- **Heads.** (1) *dwp*: salida lineal de 3 neuronas para $\hat{\Delta W P}_{\text{go}}, \hat{\Delta W P}_{\text{fg}}, \hat{\Delta W P}_{\text{punt}}$; (2) *aux_fg*: sigmoide para $p(\text{FG make})$; (3) *aux_go*: sigmoide para $p(\text{4th conv})$. Las cabezas auxiliares ayudan a establecer la distribución y los límites en las probabilidades de conversión.

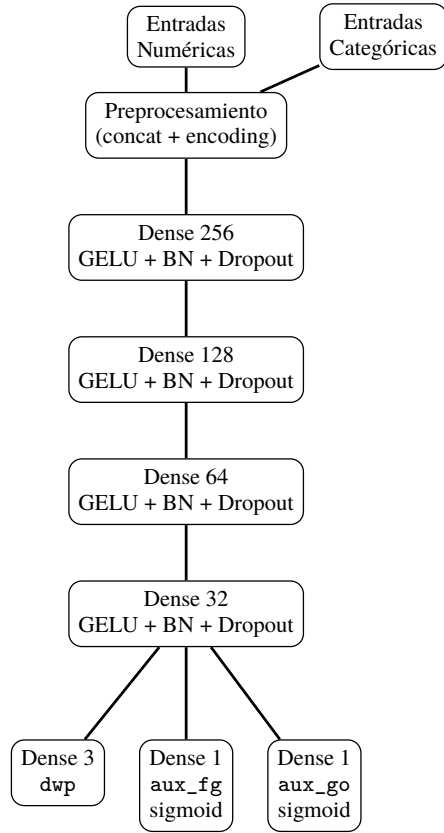


Figure 7: Esquema de la arquitectura mlp_dwp_plus_aux.

Variables y preprocesamiento

- **Numéricas (normalizadas):** quarter, game_seconds_remaining, half_seconds_remaining, yardline_100, ydstogo, score_differential, timeouts_off, timeouts_def, temp_f, wind_mph, spread_line, receive_2h_ko, fg_dist_yd.
- **Situacionales (calculadas):** in_fg_range (≤ 60 yd), one_score ($|\text{diferencial}| \leq 8$), late_game (< 600 s), second_half (cuarto > 2), timeouts_total, timeouts_diff.
- **Catóricas (one-hot):** roof $\in \{\text{open, closed}\}$, surface $\in \{\text{grass, turf}\}$.

Función de pérdida y entrenamiento

Para la cabeza principal se utilizó pérdida Huber con $\delta = 0.01$ sobre cada componente de ΔWP :

$$\mathcal{L}_\delta(e) = \begin{cases} \frac{1}{2}e^2, & |e| \leq \delta, \\ \delta(|e| - \frac{\delta}{2}), & \text{en otro caso.} \end{cases}$$

La pérdida total fue

$$\mathcal{L} = \mathcal{L}_{\text{dwp}} + \lambda_{\text{fg}} \cdot \text{MSE}(\hat{p}_{\text{fg}}, p_{\text{fg}}) + \lambda_{\text{go}} \cdot \text{MSE}(\hat{p}_{\text{go}}, p_{\text{go}}), \quad (3)$$

con $\lambda_{\text{fg}} = \lambda_{\text{go}} = 0.2$. Se empleó AdamW (LR 10^{-3} , weight decay 10^{-5}), ReduceLROnPlateau y EarlyStopping sobre val_dwp_mae (modo “min”, paciencia 6), batch 256 y hasta 60 épocas, restaurando los mejores pesos.

Método de evaluación

La partición temporal evitó fuga de información: entrenamiento (2016–2021), validación (2022) y prueba (2023–2024). Se reportan:

- **MAE** de ΔWP por cabeza (validación y prueba).
- **Policy agreement:** $\frac{1}{n} \sum_i \mathbb{1}[\arg \max_a \hat{\Delta WP}_{i,a} = \arg \max_a \Delta WP_{i,a}]$.
- **Mean regret:** $\frac{1}{n} \sum_i (\max_a \Delta WP_{i,a} - \Delta WP_{i, \pi_{\text{modelo}}(x_i)})$.
- **Análisis por bins** de posición de campo (0–20, 21–40, 41–60, 61–80, 81–100) y yards-to-go (1, 2, 3–6, 7–10, > 10), y **escenario late & close** (restan < 600 s y $|\text{marcador}| \leq 8$). Cuando se indica, se incluyen intervalos de confianza al 95% vía *bootstrap*.

RESULTADOS

Resumen global

En el conjunto de prueba (temporadas 2023–2024), el modelo alcanzó **0.794** de *policy agreement* y **0.0023** en *mean regret* de ΔWP (**0.23 %**). Estos valores indican que, en promedio, la acción sugerida por el modelo coincide con la acción óptima del *teacher* en $\sim 79\%$ de las 4^{as} y, cuando no coincide, la pérdida esperada en probabilidad de victoria es inferior a un punto porcentual.

Desempeño por región del campo y por distancia a convertir

La Figura 8 muestra el acuerdo por *bins* de posición de campo: **73.9%** (0–20), **75.1%** (21–40), **77.8%** (41–60), **84.8%** (61–80) y **85.9%** (81–100). El patrón creciente sugiere que, lejos de la propia endzone, la estructura de la decisión se vuelve más separable para el modelo (domina *punt* y los casos ambiguos disminuyen).

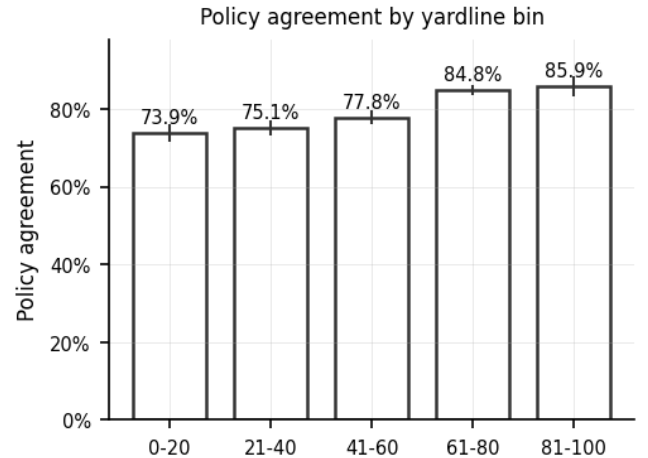


Figure 8: **Policy agreement** por *bins* de posición de campo. El desempeño mejora hacia campo propio profundo, donde *punt* domina y la frontera de decisión es más nítida.

En términos de pérdida, la Figura 9 reporta el *regret* medio (%) por yards-to-go: **0.12** (1), **0.33** (2), **0.39** (3–6), **0.28** (7–10) y **0.09** (> 10). El pico en 3–6 yardas coincide con la

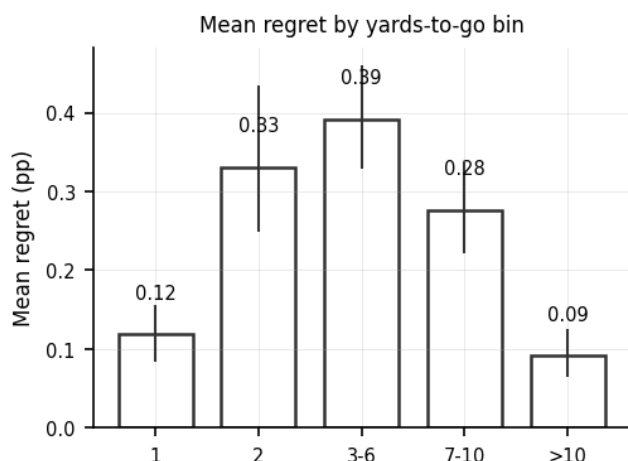


Figure 9: **Mean regret** (%) por *yards-to-go*. El máximo en 3–6 yardas refleja la zona más ambigua de decisión. Barras de error: IC 95% (bootstrap).

zona más competitiva (*trade-off* entre *go* y *punt/fg*), lo que refuerza la utilidad del asistente en ese régimen.

Errores de decisión: matriz de confusión

La Figura 10 presenta la matriz de confusión (filas normalizadas) entre la mejor acción del *teacher* y la elección del modelo. Las precisiones por fila son: **go** = 0.71, **fg** = 0.79 y **punt** = 0.88. El límite más difícil es **go** vs **punt** (confusiones del 0.16 y 0.10, respectivamente), mientras que **fg** es la acción mejor separada (sólo 0.02 se confunde con *punt*). Este patrón es consistente con los análisis por bins y con la intuición táctica.

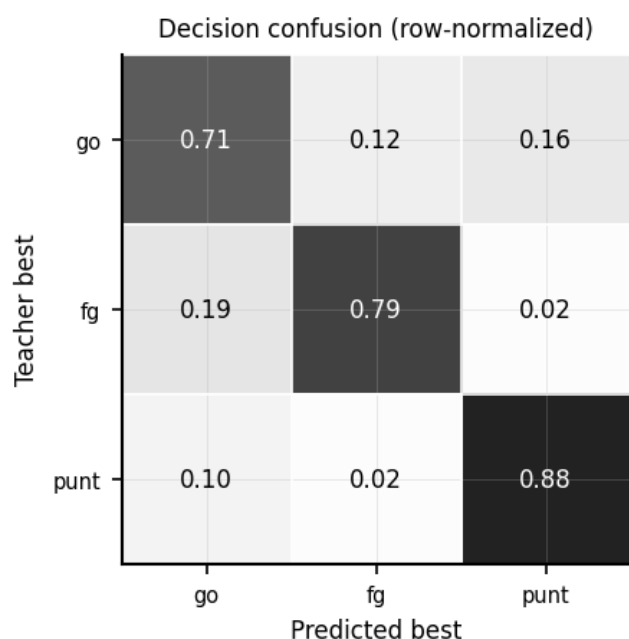


Figure 10: Matriz de confusión (filas = *teacher*, columnas = modelo). Mayor confusión en el borde *go*↔*punt*; *fg* resulta más separable.

Escenarios late & close (juegos cerrados)

En jugadas con < 600 s restantes y $|diferencia| \leq 8$, el acuerdo es **0.744** y el *regret* medio asciende a **0.0099** (**0.99** %); ver Figuras 11 y 12. El incremento de *regret* en este subconjunto es esperable: la sensibilidad de WP a la acción crece y las opciones *gol/punt/fg* se vuelven más delicadas, ya que están en momentos definitivos en un partido. Aun así, el modelo mantiene un nivel de acuerdo robusto (>70%), lo que respalda su utilidad práctica en los momentos de mayor importancia dentro de un partido.

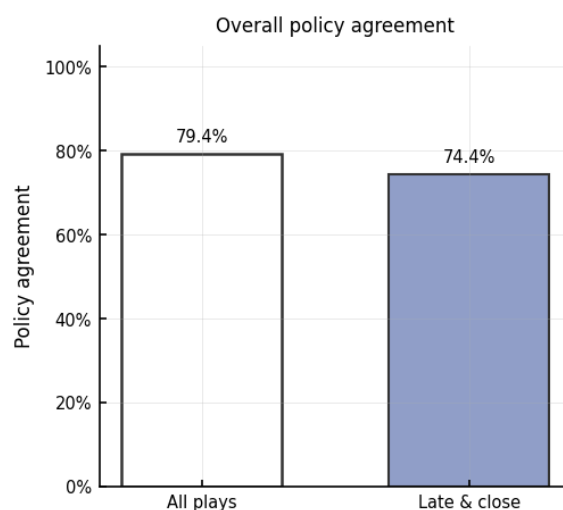


Figure 11: Acuerdo global vs. *late & close*.

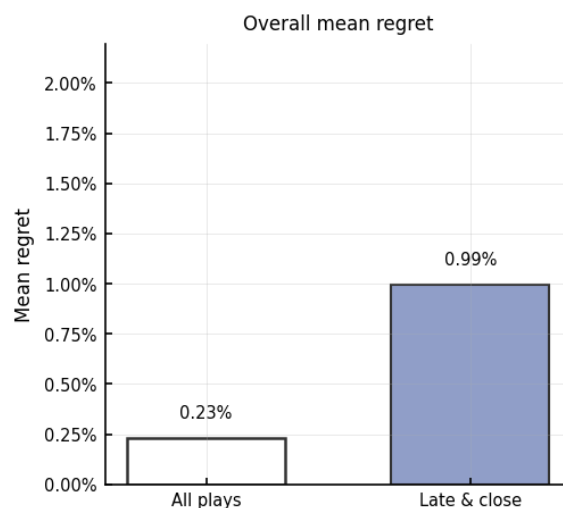


Figure 12: *Regret* medio global vs. *late & close*.

Interpretación. (i) El modelo reproduce bien la estructura de decisión del *teacher* en la mayor parte del espacio de estados; (ii) los mayores desafíos se concentran en medio campo y 3–6 yardas, donde la señal entre acciones compite; (iii) en *late & close*, la penalización por desacierto crece (casi 1%), lo que subraya la relevancia de reportar *margen de decisión* y *probabilidades similares*.

CONCLUSIONES

Este trabajo aborda la decisión de 4ª oportunidad mediante un enfoque directo de aprendizaje de ΔWP por acción. El modelo estima, para cada estado x , el vector $\hat{y}(x) = [\hat{\Delta WP}_{go}, \hat{\Delta WP}_{fg}, \hat{\Delta WP}_{punt}]$ y recomienda mediante $\arg \max_a \hat{\Delta WP}_a(x)$. De esta forma, el problema se resuelve como una comparación del efecto de cada una de las opciones *go-fg-punt* sobre la probabilidad de victoria.

Sobre datos de 2016–2024, el MLP multi-tarea consiguió en prueba **0.794** de *policy agreement* y **0.23 pp** de *mean regret*, y en *late & close* **0.744/0.99 pp**. El análisis por bins muestra que el modelo reproduce bien la estructura del *teacher*, con especial utilidad en medio campo y 3–6 yardas, donde la decisión es más competitiva.

Además, el modelo es capaz de mantener su precisión aún en momentos donde el partido está cerrado, momentos en los que un acierto o un error determina el resultado final. Esta capacidad de mantenerse en momentos importantes añade valor al modelo como una herramienta más para tomar decisiones durante el partido y también para análisis pos-partido.

REPRODUCIBILIDAD

Código y estructura. El modelo se encuentra disponible en el repositorio de github: github.com/marco-vgn/nfl-4th-down-dwp y sigue una plantilla de ciencia de datos reproducible (CCDS):

- `data/`: raw, interim, processed.
- `nfl_4thdown_dwp/`: `features.py` (extracción/limpieza de estados de 4ª).
- `labels/`: `build_labels.R` (etiquetas $wp_{go}, wp_{fg}, wp_{punt}$ vía `nfl4th` y cálculo de ΔWP).
- `notebooks/`: EDA y entrenamiento.
- `models/`: `.keras` y predicciones; `reports/figures/`: figuras.

Entorno. Python 3.11 con `tensorflow`, `pandas`, `numpy`, y R (≥ 4.2) con `nfl4th` y `tidyverse`. Las versiones exactas constan en `requirements.txt` y, para R, en el script de instalación.

Datos. Se emplean jugadas pbp públicas (2016–2024). La canalización es:

```
# 1) Extraer estados de 4ª y guardar CSV intermedio
python -m nfl_4thdown_dwp.features
```

```
# 2) Generar etiquetas con nfl4th y WP (CSV final)
Rscript nfl_4thdown_dwp/labels/build_labels.R
```

Entrenamiento. El modelo se entrena con partición temporal (train: 2016–2021; val: 2022; test: 2023–2024). Se fijan semillas para reproducibilidad (NumPy/TensorFlow). El artefacto final se guarda como:

```
models/mlp_dwp_plus_aux_best.keras
models/preds_test.parquet
```

Figuras y resultados. Todas las figuras del EDA y de resultados se generan desde los notebooks y se exportan a `reports/figures/`. Las métricas reportadas (acuerdo y *regret* globales, por bins y en *late & close*) se recalculan automáticamente al ejecutar la notebook.