



INSTITUTO POLITÉCNICO NACIONAL
ESCUELA SUPERIOR DE CÓMPUTO
(ESCOM)



Grupo: 5BM1

Materia: Algoritmos Bioinspirados

Profesora: Abril Valeria Uriarte Arcia

Proyecto Final

Alumnos:

Cornejo Morales Paola

Hernández Martínez Ernesto Ulises

Fecha de entrega:

06 de enero de 2026

Planteamiento del problema

Las apuestas deportivas de tipo parlay en la Liga MX representan un desafío complejo de optimización donde los apostadores deben seleccionar combinaciones de partidos que maximicen simultáneamente la probabilidad de acierto y el retorno económico. La dificultad radica en que, para obtener ganancia en un parlay, todos los resultados pronosticados deben ser correctos, lo que incrementa exponencialmente la complejidad del problema conforme aumenta el número de partidos seleccionados.

Actualmente, los métodos tradicionales de selección se basan en análisis intuitivos o superficiales que no integran eficientemente las múltiples variables relevantes disponibles, tales como los momios ofrecidos por las casas de apuestas, el historial de desempeño de los equipos, las estadísticas ofensivas y defensivas, y las probabilidades estimadas. Esta desconexión resulta en bajas tasas de acierto y suboptimización de ganancias, ya que no se considera adecuadamente la interdependencia entre los múltiples partidos de una jornada ni se explora eficazmente el vasto espacio combinatorio de posibles apuestas. La naturaleza multiobjetivo del problema, donde se debe equilibrar riesgo y recompensa, junto con la complejidad del espacio de búsqueda no lineal, hacen necesario el desarrollo de estrategias de optimización más sofisticadas que permitan identificar combinaciones óptimas mediante la exploración y explotación inteligente del espacio de soluciones, integrando de manera sistemática todas las fuentes de información disponibles para maximizar tanto la probabilidad de éxito como la ganancia esperada en las apuestas parlay de la Liga MX.

Algoritmo utilizado

Para la resolución del problema planteado se seleccionó el Algoritmo de la Colonia de Abejas Artificiales (Artificial Bee Colony, ABC) debido a su efectividad en problemas de optimización combinatoria con espacios de búsqueda grandes, discretos y no lineales. En este caso, la selección óptima de parlays implica evaluar múltiples combinaciones de partidos y resultados posibles, considerando simultáneamente probabilidades reales de acierto y momios acumulados, lo que da lugar a una función objetivo no convexa y con múltiples óptimos locales. El algoritmo ABC resulta adecuado al no requerir información derivada de la función de fitness y al incorporar de manera natural un balance entre exploración y explotación mediante la interacción de abejas obreras, observadoras y exploradoras. Esta característica permite evitar el estancamiento en soluciones subóptimas y favorece la búsqueda de combinaciones con mayor valor esperado. Además, su estructura poblacional facilita la evaluación simultánea de distintas soluciones candidatas, tolera la incertidumbre inherente al pronóstico deportivo y presenta una implementación sencilla con pocos parámetros, lo que lo convierte en una alternativa apropiada para el problema abordado.

Utilización dentro del problema

El Algoritmo de la Colonia de Abejas Artificiales se implementó modelando cada solución candidata como una abeja, la cual representa un parley posible dentro de una jornada de la Liga MX. Cada abeja se codificó como un vector de longitud variable n , donde n corresponde al número de partidos de la jornada, y cada componente puede tomar los valores $-1, 0, 1$ o 2 , indicando respectivamente que el partido no se incluye en el parley, o que se apuesta por victoria local, empate o victoria visitante. Las probabilidades asociadas a cada resultado se obtuvieron como el promedio entre probabilidades históricas normalizadas y las generadas por un modelo predictivo previamente

entrenado, mientras que los momios normalizados provienen de los datos oficiales de apuestas.

La función de fitness se diseñó para evaluar la calidad de cada parley propuesto por el algoritmo, considerando simultáneamente el potencial de ganancia y el riesgo asociado a la apuesta. Para ello, se definió el fitness como el producto entre el momio combinado del parley y la probabilidad conjunta de acertar todos los partidos seleccionados. El momio combinado se obtiene multiplicando los momios individuales correspondientes a cada resultado elegido (local, empate o visitante), lo cual representa el factor total de pago ofrecido por la casa de apuestas. Por otro lado, la probabilidad de acierto del parley se calcula como el producto de las probabilidades promedio de cada resultado seleccionado, donde dichas probabilidades provienen de la combinación de información histórica y las predicciones generadas por el modelo de aprendizaje automático. Esta formulación permite penalizar automáticamente aquellas soluciones con momios elevados, pero probabilidades muy bajas, así como aquellas con alta probabilidad, pero retornos poco atractivos, favoreciendo así soluciones con un mejor balance entre riesgo y beneficio.

El enjambre se dividió en abejas obreras y observadoras, donde las primeras exploran el espacio de soluciones mediante modificaciones locales, y las segundas seleccionan soluciones prometedoras utilizando un mecanismo de ruleta proporcional al fitness. Adicionalmente, se incorporó un límite de intentos para activar abejas exploradoras que reinicializan soluciones estancadas ($limit = (beehive_size * n) // 2$), así como una restricción que asegura un número mínimo de partidos por parley ($partidos_min_jornada = 2$), garantizando la validez de las soluciones generadas.

Se encontró que la combinación seleccionada de constantes para el algoritmo de colonia de abejas permitió alcanzar un balance adecuado entre la calidad de las soluciones obtenidas y el costo computacional del método. En particular, el tamaño del enjambre ($beehive_size = 100$) y el número de iteraciones ($max_iterations = 5000$) se definieron de forma que el algoritmo pudiera explorar un espacio de soluciones suficientemente amplio sin incrementar de manera excesiva el tiempo de ejecución. Asimismo, se evitó que el enjambre inicial fuera innecesariamente grande, ya que el algoritmo se ejecuta de manera independiente para cada jornada de la liga en la que es posible realizar apuestas. Considerando que el sistema debe procesar alrededor de veinte jornadas a lo largo de la temporada, el uso de un número elevado de abejas por jornada habría resultado en un incremento significativo en la carga computacional. Por ello, los valores elegidos para las constantes del algoritmo permiten mantener una ejecución eficiente, escalable y adecuada para el análisis completo de la temporada, sin comprometer la capacidad de optimización del modelo.