

---

**UTILIZACIÓN DE NAIVE BAYES PARA PREDICCIÓN DE VICTORIAS EN  
JUGADORES DE LA ASOCIACIÓN DE TENISTAS PROFESIONALES**

**MARIO RICARDO SANHUEZA LEIVA  
INGENIERÍA CIVIL EN COMPUTACIÓN**

**RESUMEN**

La predicción de juegos deportivos es un tema de interés para el mundo de las apuestas, ya sea para beneficiar a los apostadores o maximizar las ganancias de las casas de apuestas. Las distintas asociaciones de deportes han almacenado información detallada respecto al desempeño de jugadores, incluyendo resultados de enfrentamientos contra adversarios y detalles del desempeño individual, etc. Frecuentemente, mucha de esta información se encuentra públicamente disponible en los respectivos sitios web de las asociaciones deportivas. El problema a resolver consiste en utilizar esta información histórica para predecir el resultado final de un encuentro de tenis entre dos oponentes. Para lograr el objetivo propuesto, se obtuvieron datos disponibles públicamente desde el sitio web de la Asociación de Tenistas Profesionales (ATP)<sup>1</sup>. Estos datos contienen información estadística respecto a encuentros individuales entre dos oponentes e información cuantitativa del desempeño de un jugador durante una temporada. Estos datos fueron debidamente transformados y posteriormente utilizados para construir un modelo de clasificación. Se entrenó el modelo con datos de la temporada 2012, identificando que es posible predecir dependiendo del contexto, hasta con un 80% de efectividad, quién obtendrá la victoria en juegos de una misma temporada. Adicionalmente, se detectó que la superficie en la que se juega, es determinante en el proceso de predicción. Otro conjunto de experimento, consideró utilizar el modelo de aprendizaje generado a partir de datos de la temporada 2012 para predecir las temporadas 2013 y 2014. En este caso obtuvimos precisiones del orden del 65 %, lo que es superior a una predicción puramente aleatoria. En opinión de este autor, estos resultados hacen evidente el potencial de la minería de datos para anticipar resultados de eventos deportivos. Este caso se enfoca particularmente en el tenis, pero se estima que la metodología expuesta en el presente documento es suficientemente genérica para ser aplicada en otros contextos