****

Taller de proyectos 1

Sistema inteligente para pronóstico de partidos de fútbol

Project Charter

Versión 1.6

Preparado por:

|  |  |
| --- | --- |
| Jefe de Proyecto: | Michael Roland Martínez Meneses |
| Especialidad: | Ingeniería de Software |
| Empresa: | IT Expert |

Historial de Revisiones

|  |  |  |  |
| --- | --- | --- | --- |
| Versión | Fecha | Autor | Descripción |
| 1.0 | 30/10/2011 | Martínez Meneses, Michael Roland | Creación del documento |
| 1.1 | 07/11/2011 | Martínez Meneses, Michael Roland | Modificación del documento |
| 1.2 | 23/02/2012 | Martínez Meneses, Michael Roland | Modificación del documento |
| 1.3 | 19/03/2012 | Martínez Meneses, Michael Roland | Modificación del documento |
| 1.4 | 26/03/2012 | Martínez Meneses, Michael Roland | Modificación del documento |
| 1.5 | 04/04/2012 | Martínez Meneses, Michael Roland | Modificación del documento |
| 1.6 | 09/04/2012 | Martínez Meneses, Michael Roland | Modificación del documento |

Índice de Contenidos

[Resumen ejecutivo 4](#_Toc321928480)

[Marco Teórico 5](#_Toc321928481)

[Posicionamiento 7](#_Toc321928482)

[Planteamiento del problema 7](#_Toc321928483)

[Objetivos 7](#_Toc321928484)

[Objetivo General 7](#_Toc321928485)

[Objetivo Específicos 7](#_Toc321928486)

[INdicadores de éxito 9](#_Toc321928487)

[Alcance del proyecto 9](#_Toc321928488)

[El Alcance del proyecto incluirá: 9](#_Toc321928489)

[El Alcance del proyecto NO incluirá: 10](#_Toc321928490)

[Suposiciones y Restricciones 10](#_Toc321928491)

[Organización DEl Proyecto 11](#_Toc321928492)

[Equipo del Proyecto 11](#_Toc321928493)

[Stakeholders y Usuarios 12](#_Toc321928494)

[Recursos requeridos 12](#_Toc321928495)

[Fases e hitos del proyecto 13](#_Toc321928496)

[Metodologías y Enfoques del trabajo 14](#_Toc321928497)

[Riesgos y Mitigación 15](#_Toc321928498)

[Glosario de términos 16](#_Toc321928499)

[Bibliografía 18](#_Toc321928500)

[Aprobación 19](#_Toc321928501)

[Anexos 20](#_Toc321928502)

# Resumen ejecutivo

Se está proponiendo el desarrollo de una aplicación inteligente para ayudar a la predicción en las apuesta en el fútbol. De manera que pueda mejorar los ingresos de los apostadores de fútbol y además apoyar la fidelización en esta actividad.

El sistema está dirigido fundamentalmente a los apostadores de Perú que realizan sus apuestas en los portales de Betsson e Intralot, Ganagol y Te apuesto y que a través de la aplicación podrán acceder a los pronósticos de los partidos de la liga española.

Se ha considerado que los factores claves de éxito de la propuesta presentada radican en ser la primera aplicación en el Perú que cuente con un sistema inteligente para realizar buenos pronósticos de la liga española, y el uso de un algoritmo de aprendizaje que permitirá al sistema aprender y corregir sus propios errores.

Finalmente, el sistema propuesto será desarrollado a lo largo de dos ciclos regulares y la base de datos histórica se alimentará con data real. Para tal fin se requerirá de dos colaboradores de desarrollo así como dos ordenadores con el lenguaje de programación en el que se implementará la aplicación.

# Marco Teórico

De acuerdo a investigaciones realizadas por la FIFA (Federación Internacional de Fútbol Asociado) en el 2006[[1]](#footnote-1), existían 265 millones de futbolistas entre hombres y mujeres alrededor de todo el mundo, ello significó un aumento del 10% de acuerdo al mismo estudio que se realizó en el año 2000 gracias al apoyo de los 207 miembros asociados con los que cuenta dicha entidad. Ello demuestra que el fútbol es el deporte más popular en todo el mundo, debido a que existe un gran número de personas que están activamente vinculadas al fútbol. Es importante mencionar ello porque se busca mantener a dicha población vinculada al fútbol a través de las apuestas deportivas.

Las apuestas deportivas son empleadas por el ser humano como método de solvencia para satisfacer sus necesidades. Las mismas, en especial las relacionadas al fútbol, han adquirido tal protagonismo que de acuerdo a las cifras que maneja el Comité Olímpico Internacional, tal actividad mueve alrededor de 350,000 millones de dólares anuales en todo el mundo[[2]](#footnote-2). Ello, sin duda, demuestra la necesidad del ser humano en realizar apuestas en el fútbol.

Las apuestas de fútbol no han sido contrarias al Perú. Es así que la casa de apuestas Betsson (Betsson Malta Ltd) la que cuenta con más de 4000 apuestas deportivas cada día dentro de las cuales se encuentra el fútbol con ligas de todo el mundo e Intralot (Intralot del Perú S.A.C.) la cual dispone de distintos juegos de apuestas deportivas tales como Ganagol y Te Apuesto se han posicionado en el ámbito local. Todo ello ha sido posible gracias al mercado potencial que disponen en el país. Ello demuestra que la población peruana es aficionada al fútbol a tal punto que realizan apuestas en ella.

Sin embargo, la problemática a la que apunta el presente documento es dar solución a la poca eficacia del ser humano al momento de realizar sus apuestas en el fútbol. Ya que los riesgos de la poca contundencia acabarían por alejar al apostante de tal actividad o en el peor de los casos terminar por endeudar al apostante.

La mejor solución que se ha encontrado al fin descrito, radica en los logros obtenidos en un curso previo. A mediados del ciclo académico 2011 – I, el profesor David Mauricio Sánchez dejó como trabajo final del curso Sistemas Inteligentes la elaboración de un sistema experto, fue entonces que un grupo de alumnos tomó la iniciativa de elaborar un sistema experto basado en redes neuronales para realizar pronósticos de resultados de partidos de fútbol. Los resultados que se obtuvieron fueron alentadores ya que se logró predecir correctamente al menos 70% de los partidos con los cuales se hicieron las pruebas respectivas.

Es así que se desarrollará un sistema inteligente que permitirá mejorar los resultados en la predicción en base a entrenamientos. Un sistema inteligente posee un algoritmo de aprendizaje cuya función es reducir el error en la predicción mediante los sucesivos entrenamientos efectuados al sistema. Además el sistema empleará una red neuronal debido a que dicha característica permite simular el funcionamiento de las neuronas de un ser humano. En resumen, desarrollar una solución mediante un sistema inteligente permitirá simular la inteligencia humana de aprender y corregir sus errores sumándose además las ventajas de que no envejece y no se deja influenciar por emociones, características propias del hombre. Todo ello con el objetivo de que el apostante obtenga mejores resultados y que se mantenga en actividad en el mundo de las apuestas.

# Posicionamiento

## Planteamiento del problema

El problema que motiva al desarrollo del presente proyecto es la poca eficacia que tiene un apostador de eventos deportivos en el rubro del fútbol.

|  |  |
| --- | --- |
| Problema | Causas |
| Influencia de emociones  No es experto, solo un aficionado | El ser humano siente; tal condición implica muchas veces que no tome la decisión correcta.  Muchas veces el apostador no tiene los conocimientos necesarios para elegir adecuadamente y cree que solo por el simple hecho de gustarle un deporte puede predecir un resultado correctamente. |
|  |  |
| Variabilidad de los eventos deportivos  Variables que influyen  Comportamiento de los resultados | Todo evento deportivo posee de comportamientos distintos que imposibilitan tratar a un deporte como otros.  Existe una gran diversidad de variables que influyen en el resultado final de un partido de fútbol como estado anímico del jugador, actuación del réferi, entre otras que son muy difíciles de medir.  El apostador desconoce la existencia de patrones que aproximan el resultado final de un partido. |

## Objetivos

### Objetivo General

Implementar un sistema inteligente basado en redes neuronales que permita pronosticar el resultado de un partido de fútbol

### Objetivo Específicos

1. Obtener al menos 75% de aciertos para los partidos que comprenden el rango de pruebas
2. Realizar el pronóstico de un partido de fútbol
3. Emplear un algoritmo de aprendizaje supervisado
4. Modelar la predicción en base a una red neuronal
5. Implementar una red neuronal
6. Implementar un módulo de entrenamiento
7. Gestionar la información de los clubes de las ligas que comprenderán el rango de pruebas
8. Gestionar la información de los jugadores o plantillas de los clubes de las ligas que formarán parte del rango de pruebas
9. El desarrollo del proyecto debe llevarse a cabo en dos ciclos regulares

## INdicadores de éxito

A continuación se listan los indicadores de éxito del proyecto:

* El error en la predicción de los partidos de fútbol que conforman el rango de pruebas debe ser no más del 25% cada vez que haya un entrenamiento.
* Debe haberse corrido una cantidad no menos de 350 casos de prueba con datos que no hayan sido probados en la fase de entrenamiento y que la misma supere el índice de aceptación.
* Certificado de calidad entregado por parte de la empresa virtual QA y/o por quienes designe el comité de proyecto.
* Entrega del producto software con las funcionalidades especificadas en el tiempo acordado.
* Desarrollo del proyecto en dos ciclos académicos regulares.
* El proyecto “Sistema Inteligente para Pronóstico de Partidos de Fútbol” se encuentre culminado y comprenda:
  + El CD del producto software.
  + La documentación relacionada al producto software.
* Manuales relacionados al producto software.

## Alcance del proyecto

### El Alcance del proyecto incluirá:

* Empleará los datos relevantes de los clubes, tales como plantilla, torneos que disputa, posición en el torneo, partidos ganados, partidos perdidos, entre otros, que componen las ligas sobre las cuales se harán las pruebas de predicción.
* Empleará los datos relevantes de los jugadores, tales como amonestaciones, talla, peso, partidos como titular, partidos como suplente empleado, edad, entre otros, de los clubes que formarán parte de las pruebas de predicción.
* Gestionará las suspensiones de los jugadores de fútbol a lo largo de una temporada.
* Gestionará los fichajes o transferencia de jugadores entre clubes a lo largo de una temporada.
* Gestionará las posiciones de los clubes a lo largo de una temporada de las ligas que comprenderán el rango de pruebas.
* Gestionará los puntos obtenidos por los clubes en los partidos previos y que reflejen la racha que llevan.
* Gestionará los goles anotados por los clubes en los partidos previos y que reflejen la racha que llevan.
* Gestionará los goles encajados o permitidos por los clubes en los partidos previos y que reflejen la racha que llevan.
* Gestionará un ranking de clubes a nivel mundial.
* Manejará los promedios de edades de los clubes así como del equipo que frecuentemente es titular.
* Elaborará los manuales de uso del sistema inteligente.

### El Alcance del proyecto NO incluirá:

* No empleará un análisis estadístico para pronosticar el resultado de un partido de fútbol.
* No empleará un análisis del estado emocional de un jugador para pronosticar el resultado de un partido de fútbol.
* No empleará un análisis del comportamiento de los réferis de un partido de fútbol.
* No se soportará ni se hará el mantenimiento del producto luego del despliegue.

## Suposiciones y Restricciones

|  |  |
| --- | --- |
| **Suposiciones** | **Descripción** |
| Uso de liga española para entrenamiento | La liga española 2010/2011 será utilizada como liga de prueba para el entrenamiento del sistema. |
| Uso de liga española para predicción  Brindar capacitación en pruebas de SE  Emplear algoritmo de aprendizaje  Horas de labores | La liga española 2011/2012 será utilizada como liga de predicción una vez culminada la etapa de entrenamiento del sistema inteligente.  El equipo de proyecto tomará responsabilidad de dar una capacitación en el tema de pruebas de sistemas inteligentes en caso sea necesario  El equipo de proyecto tomará responsabilidad de emplear una herramienta que implemente el algoritmo de aprendizaje.  El equipo de proyecto desempeñará sus labores durante las 6 horas presenciales y adicionalmente 6 horas no presenciales de acuerdo a la disponibilidad del equipo. |
|  |  |

|  |  |
| --- | --- |
| **Restricciones** | **Descripción** |
| Tecnología a emplear | El producto software se construirá con la tecnología Microsoft .Net. |
| Disponibilidad de SO  Uso de Microsoft .Net Framework 3.0 | El producto software estará disponible para Windows XP en adelante.  El producto software empleará Microsoft .Net Framework 3.0 en el desarrollo de la aplicación. |
|  |  |

# Organización DEl Proyecto

[En este acápite se incluyen todos los detalles de organización del proyecto]

## Equipo del Proyecto

|  |  |  |
| --- | --- | --- |
| **Rol** | **Miembro** | **Responsabilidades** |
| Comité de Proyecto | Jorge Cabrera  María Hilda Bermejo  Rosario Villalta  Amanda Sánchez | Personas encargadas de establecer el plan estratégico, la aprobación de las propuestas de proyectos, decidir la continuidad de los proyectos, aprobar las contrataciones, controlar el cumplimiento de las metas y supervisar la marcha de las distintas organizaciones. |
| Gerente General | Carlos Rodríguez | Tener comunicación con los gerentes generales de las empresas virtuales. |
| Gerente de Proyectos | Julio César Guillén | Velar por que el proyecto cumpla con las fechas establecidas en el planeamiento.  Velar por que se cumpla con las capacitaciones de los colaboradores, necesarias para el desarrollo del proyecto.  Realizar un eficiente seguimiento del proyecto y evaluar a los colaboradores. |
| Jefe de Proyecto | Michael Martínez | Persona encargada de velar que el Plan de Proyectos se cumpla a cabalidad y de la coordinación con la gerencia de Proyectos y colaboradores. |
| Jefe de Desarrollo | Michael Martínez | Persona encargada de asignar el desarrollo de las funcionalidades y de integrar las mismas en la solución final. |
| Asesor de Proyecto | David Mauricio | Es profesor del curso Inteligencia Artificial y es la persona encargada de asesorar en el planteamiento del proyecto. |
| Colaboradores | Miembros del Software Factory | Llevar a cabo las tareas asignadas en las fechas establecidas.  Elaborar las tareas con calidad e iniciativa. |

## Stakeholders y Usuarios

|  |  |  |
| --- | --- | --- |
| **Stakeholders** | **Función** | **Entregables** |
| **Gerente General** | Orientar en todos los aspectos relacionados a IT. |  |
| **Gerente de Proyecto** | Es la persona encargada de velar porque el proyecto cumpla con las fechas establecidas. Realizar un seguimiento del proyecto y evaluar al proyecto. | Cronograma, artefactos |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| **Usuarios** | **Necesidades** | **Entregables** |
| **Administrador del Sistema**  **Supervisor/Maestro**  **Apostante** | Es la persona encargada de gestionar los clubes, jugadores, amonestaciones, suspensiones, lesiones y demás que utilizará el sistema.  Es la persona encargada de entrenar al sistema inteligente.  Es la persona que usará la aplicación para obtener un pronóstico en un partido específico. | Funcionalidades del Administrador del Sistema en el software final.  Modulo de entrenamiento en el software final.  Módulo de pronósticos en el software final. |
|  |  |  |
|  |  |  |

## Recursos requeridos

A continuación se hace mención a los recursos requeridos para el desarrollo del sistema:

* Visual Studio 2010 (de preferencia Edición Ultimate)
* Microsoft SQL Server 2008
* SVN
* Hudson

## Fases e hitos del proyecto

|  |  |  |  |
| --- | --- | --- | --- |
| **Hito del proyecto** | **Fecha Estimada** | **Entregables incluidos** | **Prioridad** |
| Fase: Concepción  Hito: Entrega Paquete LCO | 19/03/2012 | Exploración de Herramientas  Glosario de Términos | Baja |
| Fase: Elaboración 1  Hito: Entrega Paquete LCA | 22/03/2012 | User Stories  Priorización de User Stories  Determnación de Sprints  Determinación de Backlog  Determinación de duración de Sprints  Determinación de Release  Diseño de Base de Datos  Prototipos | Alta |
| Fase: Construcción  Hito: Entrega Primera Versión del Producto a QA | 17/04/2012 | Sprint 1 | Media |
| Fase: Construcción  Hito: Entrega Segunda Versión del Producto a QA | 21/05/2012 | Sprint 2 | Media |
| Fase: Construcción  Hito: Entrega Tercera Versión del Producto a QA | 14/06/2012 | Sprint 3 | Media |
| Fase: Construcción  Hito: Entrega Cuarta Versión del Producto a QA | 06/09/2012 | Sprint 4 | Alta |
| Fase: Construcción  Hito: Entrega Quinta Versión del Producto a QA | 27/09/2012 | Sprint 5 | Media |
| Fase: Construcción  Hito: Entrega Sexta Versión del Producto a QA | 25/10/2012 | Sprint 6 | Baja |
| Fase: Entrenamiento  Hito: Entrega Módulo de entrenamiento a QA | 20/11/2012 | Módulo de Entrenamiento | Alta |
| Fase: Pruebas SE  Hito: Entrega de Validación y verificación de SE a QA | 22/11/2012 | Resultados de Pruebas de SE | Alta |
| Fase: Documentación Final  Hito: Entrega documentación | 26/11/2012 | Versión Final del Producto  Manual de Usuario | Alta |
| Fase: Transición  Hito: Despliegue del Producto | 27/11/2012 | CD con el contenido del Producto Software | Alta |
| Hito: Aceptación del Producto | 27/11/2012 |  |  |

## Metodologías y Enfoques del trabajo

El proyecto se pretende desarrollar bajo el marco que establece la metodología SCRUM, mediante planeación de sprints. En cada sprint se implementará, se harán pruebas unitarias, entre otras actividades propias a dicha metodología. Además, el proyecto estará alojado en un repositorio para control de versiones, y se realizarán integraciones continuas con el fin de hallar fallos de manera oportuna.

## Riesgos y Mitigación

| **#** | **Riesgo** | **Probabilidad** | **Impacto** | **Estrategia de mitigación** |
| --- | --- | --- | --- | --- |
| 1 | Información insuficiente para la fase de entrenamiento del sistema | Media | Alta | Se guardarán los datos necesarios de manera progresiva antes de empezar la fase de entrenamiento, de tal manera que llegado el momento se cuente con la información requerida. |
| 2 | Paro indefinido en la liga española 2011/2012 | Baja | Media | Después de la fase de entrenamiento, se podrá simular los pronósticos y compararlos cuando se reinicie dicho campeonato. |
| 3 | Falta de eficiencia del algoritmo de aprendizaje | Media | Alta | Replanteamiento del algoritmo. |
| 4 | Red neuronal mal modelada | Baja | Media | Replanteamiento de la red neuronal |
| 5 | Mala comunicación con los stakeholders durante el desarrollo del proyecto | Baja | Media | Entablar nuevas vías de comunicación. |
| 6 | Necesidad de aumentar la cantidad de colaboradores | Baja | Alta | Contar con nuevos colaboradores según acuerdo con Software Factory |
| 7 | El software no pasa las pruebas de calidad | Baja | Media | Dar solución a las observaciones levantadas por el analista de QA. |
| 8 | Uso de una nueva tecnología | Baja | Alta | Incluir colaboradores al proyecto que estén familiarizados con la nueva tecnología |

# Glosario de términos

**Algoritmo de aprendizaje**

Es una técnica cuyo objetivo es permitir a las computadoras o sistemas aprender y corregir los errores que cometen como consecuencia de su aprendizaje.

**Apuesta deportiva**

Es una actividad que intenta predecir el resultado de un evento deportivo.

**Betsson**

Es un portal web, como muchos en el mundo, que brinda miles de apuestas deportivas cada día. Se encuentra en el Perú desde el año 2008.

**Caso de prueba**

Un conjunto de datos agrupados en base a características particulares y que es empleada en la fase de entrenamiento de sistemas expertos para determinar si el resultado es satisfactorio.

**Comité Olímpico Internacional**

Es el organismo encargado de supervisar y administrar todo en lo que respecta a los Juegos Olímpicos. Además realiza investigaciones relacionadas a todo tipo de deporte.

**Entrenamiento del sistema**

Es la actividad por la cual se prepara al sistema para aprender y corregir errores en un ámbito específico.

**FIFA**

Es el acrónimo de Federación Internacional de Fútbol Asociado. Es el ente máximo del fútbol mundial.

**Ganagol**

Es un juego de apuestas del fútbol que consiste en predecir un total de 12 partidos específicos.

**Hudson**

Es una herramienta que permite realizar integraciones continuas de proyectos. Mediante su uso se encuentran fallos en la implementación para la corrección de los mismos.

**Intralot**

Es una compañía pionera a nivel mundial en proveer sistemas integrados para la administración de loterías, juegos innovadores, y en la gestión de apuestas deportivas autorizadas por el estado.

**Microsoft SQL Server 2008**

Es un sistema que gestiona las bases de datos creado por Microsoft.

**Red Neuronal**

Es la manera de representar el funcionamiento del cerebro humano mediante el uso de computadoras y todo el potencial que estos representan ante los cálculos avanzados que son capaces de resolver a un alto nivel comparado con el cerebro humano.

**SCRUM**

Es una metodología ágil para el desarrollo de productos. Scrum tiene un enfoque iterativo e incremental para optimizar la previsibilidad y controlar los riesgos.

**Sistema experto**

Es un sistema que emula el raciocinio de un experto en alguna materia, priorizando la rapidez en la respuesta y la calidad de ésta.

**Sistema Inteligente**

Es un sistema que simula la inteligencia humana.

**Sprint**

Es una iteración de un mes de duración o menos. La duración de cada Sprint se mantiene constante a lo largo de todo el esfuerzo de desarrollo.

**SVN**

Es la abreviatura de subversion. Es una herramienta que permite controlar las versiones de los artefactos o desarrollo de un sistema.

**Te Apuesto**

Es un juego de apuestas de Intralot que consiste en elegir de una lista los partidos y sus resultados con el objetivo de ganar una apuesta deportiva.

**Visual Studio 2010**

Es un IDE, entorno de desarrollo integrado por sus siglas en inglés, que permite mediante sus lenguajes de programación desarrollar soluciones de escritorio y web.

# Bibliografía

[BETSSON (2012) Casa de apuesta online (consulta: 22 de febrero) (https://www.betsson.com/start/pe/)]

[BWIN (2012) Bwin Interactive Entertainment AG (consulta: 22 de febrero) (https://www.bwin.com)]

[EL PAIS (2007) Bwin será el nuevo patrocinador del Real Madrid (consulta: 26 de octubre de 2011)

(http://www.elpais.com/articulo/deportes/Bwin/sera/nuevo/patrocinador/Real/Madrid/elpepudep/20070611elpepudep\_13/Tes)]

[FIFA (2007) FIFA Big Count 2006: 270 million people active in football (consulta: 25 de octubre de 2011) (http://www.fifa.com/mm/document/fifafacts/bcoffsurv/bigcount.statspackage\_7024.pdf)]

[GoWin! The Football Forecaste (2012) Es un software de venta para predicción de partidos de fútbol (consulta: 22 de febrero) (http://www.gowinsoftware.com/)]

[INTRALOT (2012) Intralot del Perú dispone de los juegos Te Apuesto y Ganagol entre otros (consulta: 22 de febrero) (http://www.intralot.com.pe)]

[TERRA (2011) Blatter: "El deporte no puede tolerar manipulaciones" (consulta: 26 de octubre)

(http://www.terra.com/musica/noticias/blatter\_el\_deporte\_no\_puede\_tolerar\_manipulaciones/fox1691649)]

[RosportsVip (2012) Portal de apuestas deportivas por membresía (consulta: 22 de febrero) (http://www.rosportsvip.com/)]

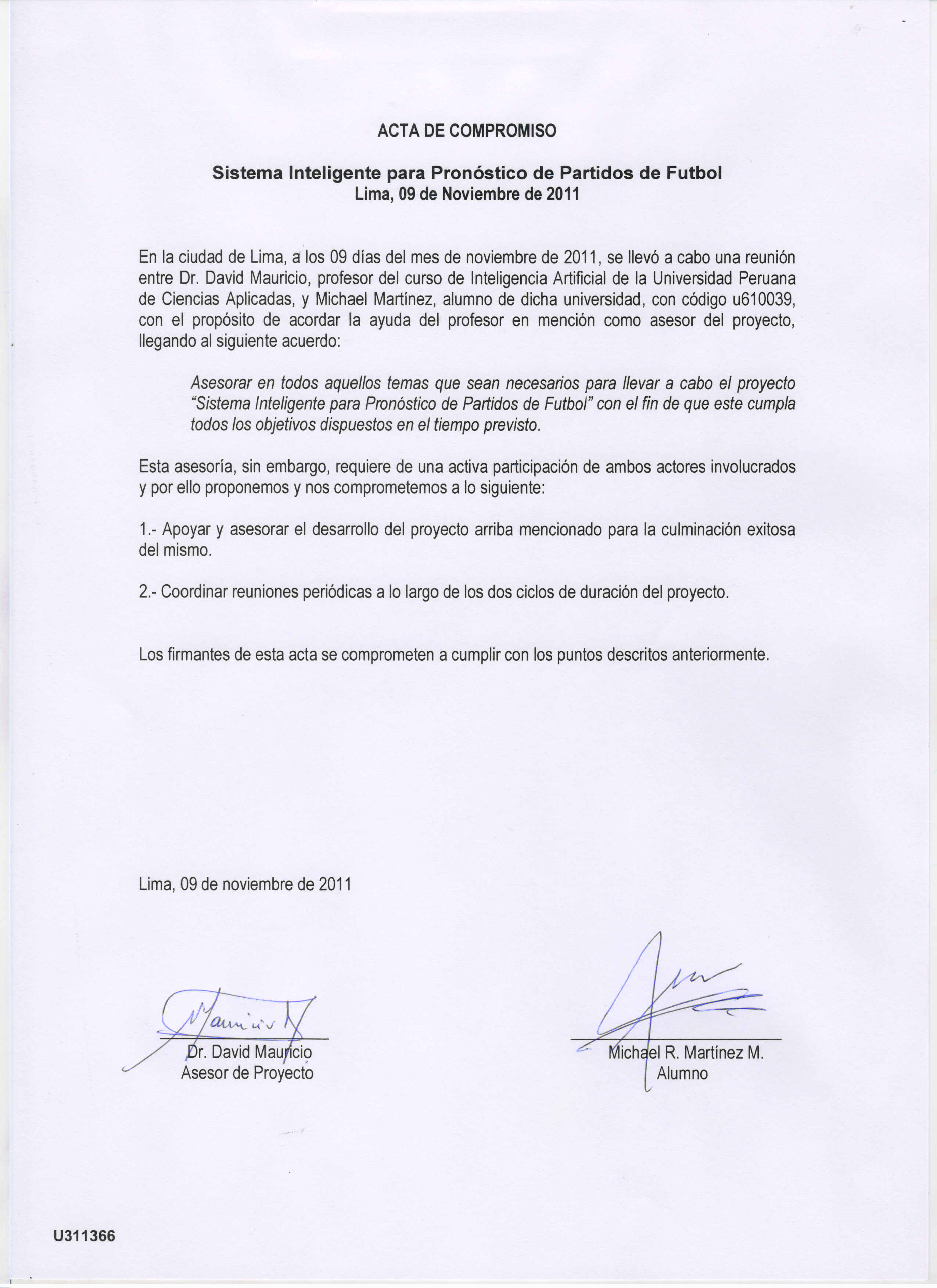
[VITIBET (2011) TOP 50 BOOKMAKERS (consulta: 27 de octubre)

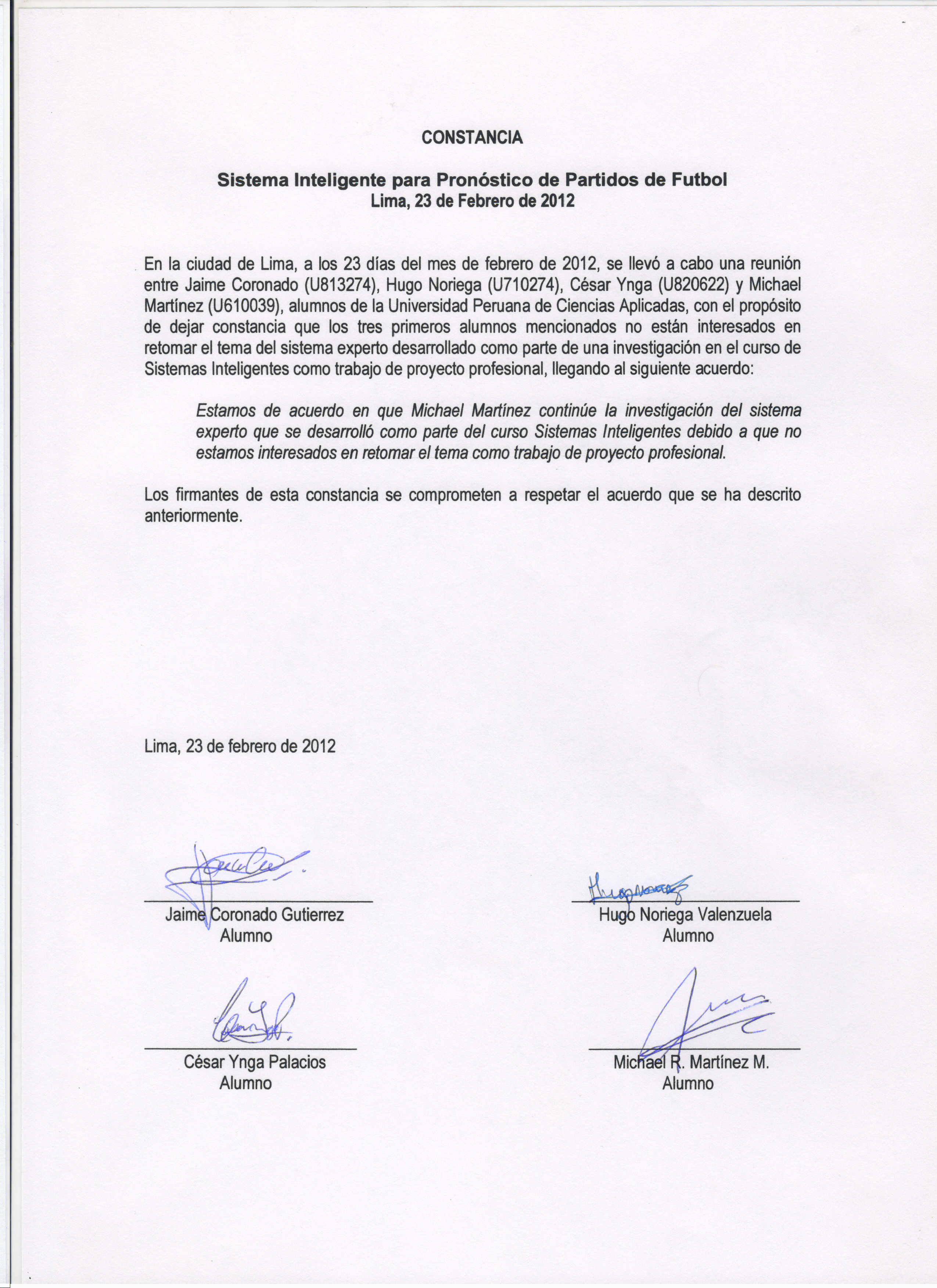
(http://www.vitibet.com/index.php?clanek=bookies&lang=en)]

# Aprobación

|  |  |  |  |
| --- | --- | --- | --- |
| **Nombre** | **Cargo** | **Firma** | **Fecha** |
|  |  |  |  |
|  |  |  |  |

# Anexos





1. http://www.fifa.com/mm/document/fifafacts/bcoffsurv/bigcount.statspackage\_7024.pdf [↑](#footnote-ref-1)
2. http://www.terra.com/musica/noticias/blatter\_el\_deporte\_no\_puede\_tolerar\_manipulaciones/fox1691649 [↑](#footnote-ref-2)