PRÁCTICA SUPERVISADA

INTRODUCCIÓN

El objetivo de este informe es documentar las actividades realizadas en carácter de alumno de práctica supervisada, en Proyecto de investigación **UTN1219 – Caracterización del Sistema Mal de Río Cuarto del Maíz mediante Minería de Datos y Análisis de Redes.** Desarrollado en el Laboratorio de Investigación de Software -Dpto. de Ing. en Sistemas de Información” en el período del 22 de Abril de 2012 hasta el 18 de Junio de 2013.

BREVE DESCRIPCIÓN DEL PROYECTO

El proyecto de Investigación **UTN1219 – Caracterización del Sistema Mal de Río Cuarto del Maíz mediante Minería de Datos y Análisis de Redes** que encuentra desarrollándose entre el 01/05/2010 y la actualidad. En el transcurso de los mismos se han generado gran cantidad de documentos (convenios, artículos, especificaciones, etc.) y buena cantidad de software en distintas plataforma, además se desarrolló un software de aplicación denominado “Yatel”.

Durante la realización del proyecto se estudia la variabilidad genética del Mal de Río Cuarto Virus a través de técnicas de Minería de Datos y el enfoque de redes.

La base de datos estará formada por muestras tomadas de individuos (plantas) enfermos y por los resultados de análisis hechos sobre estos. Los datos de las muestras incluyen, entre otros, a la ubicación geográfica del individuo, el año de la siembra, el tipo de planta (hospedante), el resultado del test de ELISA (Enzyme Linked ImmunoSorbent Assay) y el resultado del análisis de electroforésis para Mal de Río Cuarto virus (MRCV).

Las redes generadas se analizarán a través de las medidas características de las mismas y de la relación con los demás atributos de la base de datos. La relación de cada red con los otros atributos será estudiada mediante el proceso completo de Minería de Datos.

Con la aplicación de las técnicas de Data Mining (Decision Trees, Neural Networks, Cluster Analysis, Association Rules, etc) se espera lograr un mayor conocimiento del objeto en estudio, tal que permita mejorar el modelo que se tiene del mismo. Estos modelos son de gran importancia porque se utilizan para predecir la magnitud de la enfermedad en cada campaña, de cuya precisión dependen en gran medida los resultados económicos.

ROL DEL ALUMNO

Los roles desempeñados durante el período de realización de la Práctica Supervisada fueron:

* Desarrollador de Software.
* Investigador.

ANALISIS DEL PLAN DE TRABAJO

En función del plan realizado al comienzo de la práctica, se cumplimentaron la mayoría de las actividades, salvo la Preparación de presentación Final y Presentación Formal frente al equipo de Trabajo, porque hubo un pequeño desfasaje en una de las actividades. El Aprendizaje de algoritmos de redes consumió 5 horas más de lo planificado, debido a que el proyecto el mismo poseía requería la lectura de papers científicos en inglés de inteligencia artificial documentación, que debí estudiar los conocimientos de redes de grafos para la implementación de datawarehouses, empleando algoritmos de inteligencia artificial.

El desfasaje producido por las actividades mencionadas anteriormente produjo que se retrasen el inicio de las posteriores actividades y la cancelación de la presentación formal frente al grupo de investigación.

A continuación se presenta un cuadro comparativo las fechas planificadas y las fechas reales de cada actividad.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Actividad** | **Inicio planificado** | **Fin Planificado** | **Inicio Real** | **Fin Real** |
| Aprendizaje del lenguaje de Programación Python. | 22/04/2013 | 02/05/2013 | 22/04/2013 | 02/05/2013 |
| Aprendizaje del código fuente de la Aplicación “Yatel”. | 03/05/2013 | 08/05/2013 | 03/05/2013 | 08/05/2013 |
| Aprendizaje del Sistema Versionado Mercurial. | 09/05/2013 | 10/05/2013 | 09/05/2013 | 10/05/2013 |
| Aprendizaje de algoritmos de redes. | 13/05/2013 | 21/05/2013 | 13/05/2013 | 22/05/2013 |
| Programación con Python del plugin para “Yatel” implementando nuevas funcionalidades y algoritmos de redes. | 22/05/2013 | 30/05/2013 | 23/05/2013 | 31/05/2013 |
| Testing de Caja Blanca con Python de las funcionalidades y algoritmos implementados en la aplicación de escritorio “Yatel”. | 31/05/2013 | 10/06/2013 | 06/06/2013 | 11/06/2013 |
| Documentación de funcionalidades y algoritmos implementados en la aplicación “Yatel”. | 11/06/2013 | 17/06/2013 | 12/06/2013 | 18/06/2013 |
| Preparación de Presentación y Presentación del trabajo realizado en la PS a los integrantes del proyecto. | 18/06/2013 | 18/06/2013 |  |  |

DESCRIPCIÓN DEL ALCANCE DEL PROYECTO

Durante la realización del proyecto se estudiará la variabilidad genética del Mal de Río Cuarto Virus a través de técnicas de Minería de Datos y el enfoque de redes.

El proyecto requiere el aporte de desarrollo, implementación, testing y documentación empleando el lenguaje de programación Python para una aplicación de escritorio denominada “Yatel” que es empleada en proyecto en proyecto de investigación **UTN1219 -"Caracterización del Sistema Mal de Río Cuarto del Maíz mediante Minería de Datos y Análisis de Redes"**.

Yatel es una aplicación de escritorio destinada a realizar análisis de redes (grafos) para la representación de ontologías en la Minería de datos. En la actualidad el equipo de investigación del proyecto necesita añadir nuevas funcionalidades a Yatel. Dichas funcionalidades se materializan en la implementación de un plugin que involucre el desarrollo de algoritmos de análisis de redes y creación de ontologías.

Luego de completarse la fase de implementación del plugin para Yatel, es necesario realizar una fase de pruebas de caja blanca para verificar el correcto funcionamiento e integración con Yatel.

Al finalizar la fase de prueba es necesario realizar una correcta documentación tanto de la arquitectura del plugin, como de la documentación técnica necesaria para instalar dicho plugin sobre Yatel.

Finalmente al terminar el periodo de la Práctica Supervisada y cumplimentando el correspondiente desarrollo, implementación, testing y documentación pertinente al plugin para Yatel, se realizará la presentación formal a todo el equipo de trabajo, de las actividades desarrolladas.

HERRAMIENTAS A UTILIZAR

Las herramientas a utilizar para el presente proyecto son:

* Python 2.7.

Lenguaje de Programación empleado para el desarrollo de aplicación.

* Mercurial.

Sistema de versionado para mantener las versiones el código fuente de la aplicación.

* Sphinx.

Herramienta de documentación.

* Editor de Bases de Datos Relacionales.

Editor Microsoft Visio para diagramas DER.

METODOLOGÍA DE TRABAJO

En primer lugar se comenzará con el aprendizaje del lenguaje de programación Python, con el fin de posteriormente realizar la fase lectura y aprendizaje de código de la aplicación “YATEL” previamente iniciada por el Ing. Juan Bautista Cabral.

Luego se comenzará a desarrollar nuevas funcionalidades para dicha aplicación, implementando un plugin que contenga diferentes algoritmos para proveerle a la aplicación “YATEL” nuevas funcionalidades para el análisis de redes y representación de ontologías. Al finalizar el desarrollo se realizará una fase de prueba de las funcionalidades desarrolladas. Además se realizará la documentación necesaria del proyecto.

Finalizaremos con una charla en donde se presentará el trabajo realizado en la PS a los integrantes del proyecto.

EQUIPO DE TRABAJO

El equipo de trabajo está conformado por:

Director: Ing. Mario Alejandro García

Integrante: Ing. Cabral Juan Bautista

Integrante: Rodrigo Liberal

INCUMBENCIAS PROFESIONALES DEL TÍTULO DE INGENIERO EN SISTEMAS DE INFORMACIÓN

* Planificar, dirigir, ejecutar y controlar el relevamiento, análisis, diseño, desarrollo, implementación y prueba de Sistemas de Información.
* Realizar estudios e investigaciones conducentes a la creación y mejoramiento de técnicas de desarrollo de Sistemas de información y nuevas aplicaciones de la tecnología informática existente.

Materias directamente relacionadas a la Práctica Supervisada:

* Algoritmos y Estructura de Datos.
* Paradigmas de Programación.
* Análisis de Sistemas.
* Diseño de Sistemas.
* Gestión de Datos.
* Ingeniería de Software.
* Investigación Operativa.

PLAN DE TRABAJO

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Actividad** | **Inicio** | **Fin** | **Días** | **HS. Aproximadas** |
| Aprendizaje del lenguaje de Programación Python. | 22/04/2013 | 02/05/2013 | 8 | 40 |
| Aprendizaje del código fuente de la Aplicación “Yatel”. | 03/05/2013 | 08/05/2013 | 4 | 20 |
| Aprendizaje del Sistema Versionado Mercurial. | 09/05/2013 | 10/05/2013 | 2 | 10 |
| Aprendizaje de algoritmos de redes. | 13/05/2013 | 21/05/2013 | 7 | 35 |
| Programación con Python del plugin para “Yatel” implementando nuevas funcionalidades y algoritmos de redes. | 22/05/2013 | 30/05/2013 | 7 | 35 |
| Testing de Caja Blanca con Python de las funcionalidades y algoritmos implementados en la aplicación de escritorio “Yatel”. | 31/05/2013 | 10/06/2013 | 7 | 35 |
| Documentación de funcionalidades y algoritmos implementados en la aplicación “Yatel”. | 11/06/2013 | 17/06/2013 | 5 | 25 |
| Preparación de Presentación y Presentación del trabajo realizado en la PS a los integrantes del proyecto. | 18/06/2013 | 18/06/2013 | 1 | 5 |
| Total |  |  |  | 205 |