

SOLICITUDE DE APROBACIÓN DE ANTEPROXECTO DO TRABALLO DE FIN DE GRAO GRAO EN ENXENARÍA INFORMÁTICA.

Entréguese na Administración da ETSE antes da data indicada na convocatoria.

Datos do/a Alumno/a			
Nome: Pablo Pérez Romani		DNI: 78807092F	
Enderezo: Avenida da Constitución nº 27		Localidade: Portosín	
Provincia: A Coruña	C.P.: 15999	Teléfono: 662074622	Correo-e: pablo.perez.romani@gmail.com

Requisitos
O alumnado poderá ter pendentes como máximo 75 créditos para completar os estudos, excluídos os correspondentes ao TFG

Datos do Traballo de Fin de Grao	
Título: JDataMotion: unha ferramenta para a visualización dinámica de diagramas de dispersión	
Titor/a: Paulo Félix Lamas	Correo-e: paulo.felix@usc.es
Cotitor/a: David González Márquez	Correo-e: david.gonzalez.marquez@usc.es
Cotitor/a:	Correo-e:
Áreas de coñecemento: CCIA	
Departamento: Electrónica e computación	

A persoa que asina e cos datos que se indican solicita á Comisión de Traballos Fin de Grao de Enxeñaría Informática a aprobación do anteproxecto que se acompaña.

Santiago de Compostela, de de 20__

O/A alumno/a

Vº e Pr. O/A Titor(a)/Cotitor(a) do proxecto

Asdo:

Asdo:

Á atención da Comisión de Traballos Fin de Grao de Enxeñaría Informática

DESCRINCIÓN DO ANTEPROXECTO

Título

JDataMotion: unha ferramenta para a visualización dinámica de diagramas de dispersión

Introdución

Na actual sociedade da información, onde a cantidade de datos que se manexan aumenta día a día de xeito exponencial, a minería de datos convértese nunha ferramenta fundamental para poder explotalos de maneira eficaz, co fin último de xerar coñecemento a partir dos mesmos.

Para visualizar estes datos unha das técnicas máis utilizadas son os diagramas de dispersión ou scatterplots. Estes permítennos analizar os datos e atopar con facilidade relacións entre as distintas variables.

Non obstante, os scatterplots restrínxennos a unha perspectiva estática do problema. En moitos deses problemas imos encontrar unha compoñente dinámica intrínseca como é o tempo. Con este proxecto pretendemos dotar a esta representación da súa perspectiva dinámica, para amosar os datos engadindo outro punto de vista que enriqueza a información extraída.

A utilidade de unha ferramenta como esta xa foi validada no campo da electrocardiografía, pero sen ningunha dúbida haberá moitos outros ámbitos que poderán sacar proveito da análise de datos desde unha perspectiva dinámica.

Obxectivos

O obxectivo deste proxecto é o desenvolvemento dunha ferramenta capaz de visualizar a modo de “película” a evolución dun conxunto de datos ao longo dunha magnitude como sería o tempo. Sendo máis específicos, este proxecto persigue os seguintes obxectivos:

- Brindar ao usuario unha aplicación que facilite o procesado de volumes de datos dun tamaño significativo.
- Posibilidade de traballar con formatos CSV ou ARFF.
- Dispor das funcionalidades necesarias para manipular os datos.
- Ser capaz de amosar os datos en forma de scatterplots, con funcións de reprodución básicas. Tamén se debe posibilitar a configuración desta reprodución por parte do usuario.
- Aplicar filtros nos datos cos que se traballa.

Descrición técnica

A ferramenta desenvolverase integramente na linguaxe de programación Java. A súa interface gráfica estará programada utilizando a librería gráfica Swing.

A nivel funcional, JDataMotion busca estender coa perspectiva dinámica as posibilidades do software de Weka[1]. En base a isto, aproveitaremos algunhas características da súa interface ou incluso código da librería Weka para desenvolver JDataMotion.

O seu funcionamento parte dun arquivo dado en formato CSV ou ARFF que deberá importar, ou ben dunha sesión (en formato JDM) gardada durante un experimento anterior. Como ben se mencionou, botarase man das librerías facilitadas pola ferramenta Weka[1] para a importación, exportación e almacenamento do modelo de datos.

Os tipos de datos dos atributos poderán ser configurados de acordo ás seguintes etiquetas propias do estándar que proporciona o formato ARFF[3]: nominal, numérico, string ou data. Tamén se pode especificar que un atributo actúe como índice temporal para ser utilizado na reprodución. Así mesmo, permítese a inserción, eliminación e modificación dos datos.

Poderanse engadir filtros configurables aos datos que se están procesar. Estes terán efecto acumulativo, pero poderase navegar entre modelos de datos intermedios. Tamén se proporcionará unha interface para que un usuario poida aplicar no seu experimento filtros personalizados en formato JAR

De acordo coa ferramenta Weka[1] para a visualización de datos seguirase un esquema matricial, de xeito que se poida amosar nunha matriz un scatterplot para cada dupla de atributos a representar. Os scatterplots dentro da matriz estarán ordenados por filas e columnas segundo o atributo que representen en cada eixo. Para a creación de scatterplots utilizaremos unha librería de Java chamada JFreeChart[2].

A visualización terá opcións de reprodución e será configurable. Poderase sinalar un atributo nominal para que os seus valores se representen con puntos de cor e forma diferente e ampliar un scatterplot nunha ventá diferente, así como facer zoom e reposicionar a ventá de cada scatterplot.

O usuario poderá exportar o seu traballo en calquera momento baixo un novo ficheiro de formato ARFF ou CSV, ou ben gardar a sesión (JDM) para retomala máis tarde.

Medios materiais necesarios.

- Ordenador con calquera sistema operativo, e unha versión de Java preferiblemente actualizada (Java RE 1.8)

Bibliografía.

- [1] <http://www.cs.waikato.ac.nz/ml/weka/>
- [2] <http://www.jfree.org/jfreechart/>
- [3] <http://www.cs.waikato.ac.nz/ml/weka/arff.html>

Observacións.

FASES DO TRABALLO E ESTIMACIÓN TEMPORAL

Un traballo de fin de grao suporá 401,25 horas de traballo autónomo do alumnado e 11,25 horas de traballo presencial (titorías e avaliación).

Dedicación semanal prevista (en horas/semana): 15,28

Fase	Estimación temporal (en semanas)
Sprint 00	2
Lectura da especificación e documentación	0.5
Reunión cos directores do proxecto	1
Redacción do anteproxecto	0.5
Sprint 01	1
Importar arquivos	0.25
Exportar datos	0.25
Abrir sesión	0.25
Gardar sesión	0.25
Reunión co director para probas e aprobación	0
Sprint 02	2
Representar os datos en forma de táboa	0.75
Insertar datos	0.25
Modificar datos	0.5
Eliminar datos	0.5
Reunión co director para probas e aprobación	0
Sprint 03	2
Asignar tipos aos atributos	1.5
Sinalar identificación temporal	0.5
Reunión co director para probas e aprobación	0
Sprint 04	3
Representar graficamente mediante scatterplot	1
Engadir scatterplots ao menú de visualización	0.5
Eliminar scatterplots do menú de visualización	0.5
Configurar scatterplots do menú de visualización	1
Reunión co director para probas e aprobación	0
Sprint 05	2

Informar sobre punto seleccionado	0.25
Resaltar punto en todos os scatterplots	0.75
Reposicionar ventá do scatterplot	0.5
Zoom in e zoom out en scatterplot	0.5
Reunión co director para probas e aprobación	0
Sprint 06	1
Reproducir a secuencia de datos (Play)	1
Reunión co director para probas e aprobación	0
Sprint 07	2
Difuminar puntos ao longo da reprodución	0.25
Representar estela	0.5
Difuminar estela ao longo da reprodución	0.25
Configurar reprodución	1
Reunión co director para probas e aprobación	0
Sprint 08	2
Rebobinado e avance rápido (Rewind, FastForward)	0.25
Pausar a reprodución (Pause)	0.75
Ir a un determinado punto da reprodución (GoTo)	0.5
Escalar e reposicionar dinámicamente	0.5
Reunión co director para probas e aprobación	0
Sprint 09	2
Insertar filtros	1
Eliminar filtro	0.25
Configurar filtro	0.75
Reunión co director para probas e aprobación	0
Sprint 10	2
Gardar subsecuencia de filtros	0.75
Cargar subsecuencia de filtros	0.75
Mover filtros dentro da secuencia de filtros	0.5
Reunión co director para probas e aprobación	0
Sprint 11	1
Calcular distancia entre dous puntos	0.25
Configurar fórmula da distancia	0.5
Configurar menú de visualización	0.25
Reunión co director para probas e aprobación	0
Documentación	5
Revisión e documentación do código	0.75
Compilación de documentos dos sprints	0.25
Redacción da memoria	3
Redacción dun manual de usuario	1
Reunión co director para revisión	0
Total	27