

Análisis, evaluación y predicción de riesgos financieros mediante métodos de aprendizaje automático

Jornada S2B B2S Transfer N° 1

Estructura interna del INIT

Aprendizaje
automático y
reconocimiento
de patrones



Imagen y
Modelado
Matemático



Información
geográfica



Ingeniería
visual



Digitalización y
procesamiento
de textos



Digitalización y
procesamiento
de textos



Óptica



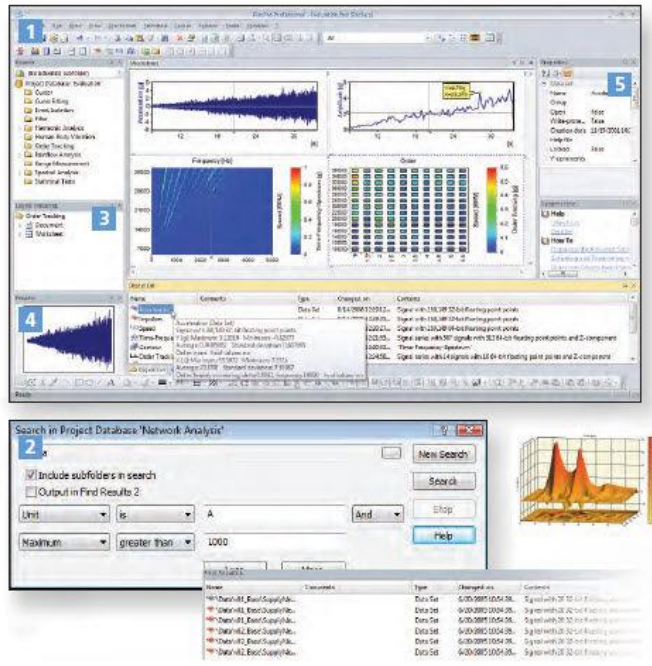
Modelado
basado en
bocetos



- Investigadores de diversos ámbitos científicos
- Laboratorios con equipamiento especializado
- Amplia experiencia en el desarrollo de proyectos de I+D+i y contratos de transferencia de resultados con empresas privadas y entidades públicas
- Cursos de especialización y postgrado

Sección Aprendizaje Automático

**Análisis, evaluación y explotación de datos:
económicos, financieros, biomédicos, comerciales,
climatológicos, demográficos, etc.**



Vídeo-vigilancia, control de accesos y verificación de identidad mediante análisis de rasgos biométricos

Qué es el Aprendizaje Automático?

Conjunto de métodos y herramientas que permiten procesar y analizar datos para extraer información útil y objetiva

- Métodos de predicción, clasificación y regresión
- Reglas de asociación
- Algoritmos de clustering
- Técnicas para el análisis de series temporales

Aprendizaje Automático en gestión financiera

- Predicción del riesgo de devolución de crédito o de impagos
- Análisis predictivo de fallida financiera o bancarrota
- Detección de operaciones fraudulentas
- Evaluación de tendencias (sentimientos) del mercado
- Segmentación de clientes: identificación de perfiles de clientes
- Fidelización de clientes: detección temprana de pérdida de confianza
- Seguimiento y análisis de patrones demográficos
- Comprobación de identidades

Tipos de datos para gestión financiera

Modelos de Big Data

- Datos estructurados (modelos clásicos de aprendizaje automático)
 - Ratios financieros
 - Datos económicos
 - Datos socio-demográficos
- Datos no estructurados
 - Datos generados en redes sociales, foros, e-mails, blogs, mensajería instantánea, formularios, ...
 - Vídeo e imágenes
 - Audio



Beneficios del Big Data

- Simplificación y mejora en la toma de decisiones (más inteligentes)
- Reducción de costes
- Mejor conocimiento del mercado y de nuestros clientes y competidores
- Fácil acceso a nuevas fuentes de datos
- (Casi)-inmediatez en el tiempo de respuesta
- Ofertas (de productos y servicios) optimizadas y personalizadas
- Anticipación a los problemas



Obtener más información y conocimiento

Cómo puede ayudar el Aprendizaje Automático?

- Aspectos diferenciales por el uso del Aprendizaje Automático en las aplicaciones de Big Data:
 - Apoyo en procesos complejos de toma de decisiones
 - Integración de datos de fuentes internas y externas
 - Explotación de grandes volúmenes de datos (Petabytes = 10^{21} bytes, Exabytes = 10^{18} bytes, Zettabytes = 10^{21} bytes) de forma más simple
 - Capacidad de combinar datos estructurados y no estructurados
 - Análisis de flujos de datos (en tiempo real)
 - Adaptación a entornos dinámicos (cambios socio-políticos, económicos, ...)

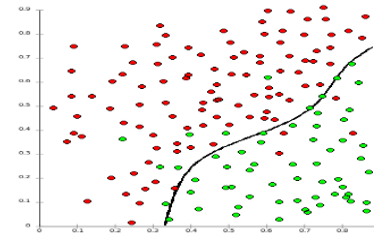
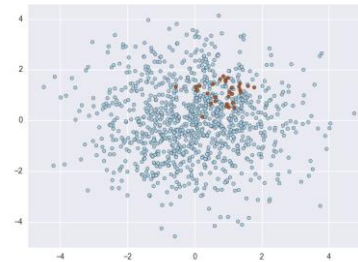
Nuestra experiencia en este campo (I)

- **Uso de los datos: evaluación y predicción del riesgo crediticio, predicción de bancarrota, predicción de fraudes**
 - **Combinación de clasificadores: bagging, boosting, rotation forest, random subspace, stacking**
 - **Redes neuronales: MLP, RBF, red bayesiana, memoria asociativa, red convolucional**
 - **Árboles de decisión: CART, C4.5**
 - **Métodos lineales: SVM, discriminante lineal de Fisher**
 - **Modelos de regresión lineal, logística y SVM**
 - **Sistemas multi-criterio para la toma de decisión: TOPSIS, PROMETHEE, ELECTRE**

Nuestra experiencia en este campo (II)

➤ Calidad de los datos: procesamiento de datos financieros

- Desbalance entre clases
(default <<< non-default)
- Detección de casos atípicos
- Selección y ponderación de variables



Nuestras bases de datos financieras

Data set	Variables	Default/non-default cases	% Default cases
Australian	14(8)	307/383	44.5
German	24(21)	300/700	30.0
Japanese	15(9)	296/357	45.3
Iranian	27(4)	50/950	5.0
Polish	30(0)	112/128	46.7
SabiSPQ	19(3)	590/590	50.0
Thomas	12(4)	323/902	26.4
UCSD	38(6)	599/1836	24.6
USA	17(0)	320/7973	3.9

Nuestros objetivos a corto plazo

- **Analizar, evaluar y explotar datos de entidades próximas**
- **Combinar datos estructurados y no estructurados para un mejor conocimiento del mercado y de los clientes**
- **Ampliar nuestra investigación a las áreas aún no exploradas:**
 - **Evaluación de tendencias de mercado**
 - **Identificación de perfiles de clientes**
 - **Seguimiento y análisis de patrones demográficos**
 - **Fidelización de clientes**