Eduardo San Miguel Martín

eduardo.san.miguel.martin@gmail.com

Summary

*Result-oriented, hands-on **data science professional**, with a successful record of accomplishments in **Machine Learning, Text Mining, Business Intelligence and Analytics** projects, as well as in academic research.*

*Over **fifteen years of professional experience** in **diversified industry fields**, covering data modeling, data mining, text mining, machine learning and NLP research & development. An experience that not only leads to enriched industry understanding, but also aids exchange of ideas, helping me build a holistic perspective.*

*Major strengths include **strong leadership** and capacity to work as a **team player**, as well as **excellent communicational and technical skills**.*

Highlights

- **Big Data Analytics**
- **Business Intelligence**
- **Data Mining & Machine Learning**
- **Text Mining & Natural Language Processing**
- **Social Media Analytics**
- **Data Warehouse Modeling**
- **Supply Chain Optimization & Warehouse Simulation**

- **R, Python, Scala**
- **SQL**
- **Oracle, Db2, SAP Hana**
- **Hbase, mongoDB**
- **Hadoop Ecosystem (Hadoop, hive, flume, sqoop)**
- **Spark & Flink**
- **Splunk**

Experience

Senior Big Data & Analytics Consultant

05 / 2016 – CURRENT

02 / 2008 – 05 /2016

- Conceive and design **end to end big data machine learning, text mining and analytics solutions** to support business units and initiatives.
- Work with business owners to map **business requirements into technical solutions**.
- Develop and implement analytic solutions, which may include applying algorithms from a standard tool or **custom algorithm development**.
- Data **warehouse modeling**.
- Support production **data mining and text mining** solutions.
- Perform ad hoc statistical and data mining analyses.
- **Reporting** of analyses to business units.
- **Visualize** and report data findings creatively in a variety of visual formats that appropriately provides **insights** to the organization.
- **Mentoring** and training other employees in **data science** and big data related skills

Retail:
- Warehouse Simulation & Supply Chain Optimization at Inditex

Financial Services:
- Financial Risk Modeling at Popular Bank (formerly Pastor Bank).
- Reporting Services for public financial statements and confidential statistical returns according to Banco de España regulatory rules at Popular Bank (formerly Pastor Bank).
- IT Control Objectives for Basel II at Popular Bank (formerly Pastor Bank).
- Fraud Detection Techniques at Spanish Tax Agency (AEAT. Agencia Española Administracion Tributaria).

Social Media Analytics:
- Social Media Monitoring & Analytics at Indra Digital.
- Topic extraction and text categorization and clustering for massive datasets at Indra Digital.
- Social Web Mining & Sentiment Analysis Model (iThink) at Indra Digital.

Cyber Security:
- Real-time Proactive and Predictive Cyberattacks Detection System: BIGUARDIAN - FEDER INNTERCONECTA 2015

Bioinformatics:
- Risk stratification and classification of chronic disease (SIXIC FEDER-INNTERCONECTA project) at SERGAS - CDTI (Centro Desarrollo Tecnológico Industrial).

Telecom & Energy:
- Electricity Consumption & Customer Segmentation at Indra.
- Targeted Churn prediction at Telefónica.
- Electric Demand Response Prediction and Customer Classification (PRICE. Joint Project of Intelligent Networks in the Henares Corridor) at Iberdrola and Gas Natural Fenosa.

Former Experience

Senior IT Consultant
GTT SL
*September 2007 –
January 2008*

*ML & Text Mining
Consultant*
Thales
*February 2006 – August
2007*

*ML & Text Mining
Consultant*
Desarrollo Humano e
Inteligencia Artificial
SL
*March 2004 – February
2006*

Education

Universidad Complutense de Madrid
Md Cognitive Sciences
Computational linguistics and Machine Learning
Certificado de Docencia e Investigación (2000-2002)
Diploma Estudios Avanzados - Suficiencia Investigadora (2002)
Certificado Defensa Publica Tesina (2003)

Instituto Madrileño de Formación
Master en Prevención de Riesgos Laborales
Especialidad en Seguridad Laboral (2003)
Especialidad en Higiene Industrial (2003)
Especialidad en Ergonomía y Psicología del Trabajo (2003)

Universidad Complutense de Madrid
Md Psychology
Cognitive & Experimental Psychology specialist $(1995 - 2000)$

Languages

Spanish: Native

English: Highly proficient in both spoken and written English

Professional Training

Cloudera Developer Training for Spark and Hadoop (CCA) 60 horas.
Cloudera University. CNTG, 2016

Cloudera Developer Training for Apache Hadoop (CCDH) 60 horas.
Cloudera University. CNTG, 2014

Herramientas Computacionales Avanzadas y Redes Neuronales Aplicadas al Análisis Estadístico en el Entorno Empresarial (94 Horas) Universidad Carlos III, 2002

Publications

BiGuardian: Sistema de detección proactiva y predictiva de amenazas de ciberseguridad.
ITMATI: Machine Learning Workshop Galicia 2016

Developing Web Apps in R with Shiny
I Xornada Usuarios R en Galicia. Centro de Novas Tecnoloxías. 2013
http://webmelisa.es/xornada-r/

fisheyeR: Fisheye and Hyperbolic-space-alike Interactive Visualization Tools in R
CRAN - The Comprehensive R Archive Network. 2010
http://cran.r-project.org/web/packages/fisheyeR/index.html

Implementación de un Modelo de Visualización para Tareas de Recuperación de Información
XXVII INTERNATIONAL AESLA CONFERENCE. 2009
Corpus linguistics, computational linguistics and linguistic engineering

Herramienta de representación tridimensional de conocimiento basada en la relación de LSA extraída a partir de corpus

Corpus linguistics, computational linguistics and linguistic engineering

3D Knowledge Retrieval