Zementis

Universal Deployment of KNIME Models

Big Data and Real-time Scoring

KNIME User Group Meeting, Berlin February 2015

www.zementis.com
Zementis provides software for operational deployment of predictive analytics

Products & Capabilities:

• Vendor-neutral architecture for
  - Data mining tools
  - Analytics and data warehouse platforms
• Supports PMML industry standard and wide range of predictive modeling techniques
• Rapidly deploys and executes predictive models
• Accelerates business insight

KNIME / Zementis Partnership

• Predictive Model Markup Language (PMML)
• Open standards & open platform
• Fosters transparency and best practices
• Instantly deploy your KNIME models everywhere
• Real-time or Big Data
What is PMML?

Predictive Model Markup Language (PMML) industry standard reduces the complexity of operationalizing models

- **Mature standard developed by the DMG** (Data Mining Group) to avoid proprietary issues and incompatibilities and to deploy models

- **XML-based language** used to define statistical and data mining models and to share these between compliant applications

- **Supported by most leading data mining tools**, commercial and open-source

- **Data handling and transformations** (pre-and post-processing) are a core component of the PMML standard

- **Allows for the clear separation of tasks**: Model development vs. model deployment

- **Eliminates the need for custom code** and proprietary model deployment solutions
One Standard, One Process

Applications

PMML

External Vendors

Divisions

Service Providers
Execute Anywhere

Vendor-neutral architecture enables compatibility with industry-leading analytics and data warehouse platforms

### Data Mining Tools
- Compatible with most leading commercial and open-source data mining tools
- Other tools include:
  - IBM SPSS
  - Python
  - R
  - SAS

### Predictive Modeling Techniques
- Supports multiple broad categories of predictive modeling techniques
- Examples:
  - Clustering Models
  - Decision Trees
  - Neural Network Models
  - Naïve Bayes Classifiers
  - Random Forest Models

### Analytics and Data Warehouse Platforms

#### Cloud
- Amazon Web Services
- FICO Analytic Cloud
- Microsoft Azure Cloud

#### On-site
- IBM WebSphere
- RedHat JBoss / Tomcat
- Oracle WebLogic
- SAP HANA

#### In-database
- IBM PureData / Netezza
- Pivotal / EMC Greenplum
- SAP Sybase IQ
- Teradata + Teradata Aster

#### Hadoop
- Hive / Storm / Spark
- Datameer
- Cloudera
- Hortonworks
- IBM
Broad Applicability

ADAPA and UPPI accelerate predictive model insights for multiple industries and business use cases

**Fraud & Risk Scoring**
- Financial institutions
- Scoring bureaus
- Online transaction processing
- Advanced decision management

**Internet of Things (IoT)**
- Predictive maintenance
- Quality control
- Sensor & device data processing
- Manufacturing
- Healthcare

**Marketing & Sales**
- Up-/cross-sell and next-best-offer
- Marketing campaign optimization
- Real-time recommendations
### Thank You!

#### USA Headquarters

3570 Carmel Mountain Road  
Suite 300  
San Diego, CA, 92130  

Phone: +1 619 330-0780

#### San Francisco Office

123 10th Street  
San Francisco, CA 94103

Phone: +1 415 800 5461

#### Asia Office

19/F., Unit A  
Ho Lee Commercial Building  
38-44 D’Aguilar Street  
Central, Hong Kong (S.A.R.)

Phone: +852 2868-0878
KNIME ▶️ PMML ▶️ Zementis

Best Practices for Predictive Analytics and Data Mining

- Open Standards vs. Proprietary Code
- Select Best-of-Breed Tool Set
- Avoid Vendor Lock-in

- Deploy in Minutes vs Months
- Facilitate Clear Requirements & Communication
- Scale with Business Demand

- Big Data & Real Time
- In-Database & Hadoop
- Server, Cloud & SaaS

Vendor Neutral Standard

Time-to-Market Agility

Platform Independent Deployment