KDD 2016 - Special Session: Standards in Predictive Analytics

Robert Grossman
Open Data Group
University of Chicago

Walt Wells
Data Mining Group

Michael Zeller Zementis
ACM SIGKDD
ACM SIGKDD Standards Initiative

• Leading the adoption of standards for the practical application of data science

• In partnership with DMG
  – Predictive Model Markup Language (PMML)
  – Portable Format for Analytics (PFA)
2016 Activities

• Roundtable “Standards in Predictive Analytics”
  – Focused on engaging the community & decision makers
  – June in Chicago
  – October in New York City
  – November in San Diego

• KDD 2016 Special Session
  • Invited talks from Industry leaders
  • Panel discussion with vendors
A brief history of the DMG

DMG Founded
PMML v0.7 released
PMML v0.9 released
PMML v1.1 released
PMML v2.1 released
PMML v3.1 released


CCSR Founded
PMML v2.0 released
PMML v3.0 released
Portable Format for Analytics (PFA) Introduced
Portable Format for Analytics (PFA) support begins


PMML v4.1 released
PMML v4.2.1 released
PMML v4.3 released
PMML v3.2 released
PMML v4.0 released
Why are standards important?

Data Scientists

- Models
  - Data mining tools
    - Commercial
    - Open-source
  - Algorithms
    - Statistical
    - Machine learning

IT Engineers

- Infrastructure
  - Cloud
  - In-house
  - In-database
  - Hadoop

Business

- Applications
  - Real-time Scoring
  - Batch Scoring

BUILD

DEPLOY

USE
How can you get involved?

• Join the DMG
• Visit the PMML discussion group on LinkedIn
• Get the flyers and learn more about standards
# Agenda

## Special Session Program

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Speaker/Panelist</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:30</td>
<td>Welcome &amp; Remarks</td>
<td>Michael Zeller, CEO - Zementis</td>
</tr>
<tr>
<td>09:40</td>
<td>Invited Talk</td>
<td>James Reid, VP Software Development - Equifax</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Write Once, Run Anywhere</td>
</tr>
<tr>
<td>10:00</td>
<td>Invited Talk</td>
<td>Soundar Srinivasan, Chief Expert - Data Mining - Robert Bosch, LLC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The Value of Data Science Standards and First Experiences in Using PMML in Manufacturing Analytics</td>
</tr>
<tr>
<td>10:20</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>10:35</td>
<td>Invited Talk</td>
<td>Joydeep Das, Chief Operating Officer - KloudData</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A Standards Based Prescription for Healthcare Analytics</td>
</tr>
<tr>
<td>10:55</td>
<td>Invited Talk</td>
<td>Paco Nathan, Director, Learning Group - O'Reilly Media</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of Standards and Related Issues in Data Analytics</td>
</tr>
<tr>
<td>12:15</td>
<td>Adjourn</td>
<td></td>
</tr>
</tbody>
</table>
Speakers

Invited Speakers

Joydeep Das
Chief Operating Officer
KloudData

Paco Nathan
Director, Learning Group
O'Reilly Media

James Reid
VP Software Development
Equifax

Soundar Srinivasan
Chief Expert - Data Mining
Robert Bosch, LLC
Panelists

Stuart Bailey
Chief Technology Officer
Open Data Group

Jim Pivarski
Associate Professional Specialist
Computational Physicist at Princeton

Shalini Raghavan
Sr. Director Product Management
FICO

Lawrence Spracklen
Vice President of Engineering
Alpine Data

Wayne Thompson
Chief Data Scientist
SAS Data Science Technologies

Michael Zeller
Chief Executive Officer
Zementis