MIKE
FERGUSON

MACHINE LEARNING

AND ADVANCED ANALYTICS

ONLINE LIVE STREAMING

DECEMBER 1-2, 2020

ABOUT THIS SEMINAR

Today, with most people connected to the Internet, the power of the customer is almost limitless. The Internet has given them freedom to choose in a way that business could never have imagined. They can browse your competitors' Web sites with ease. They can compare prices, they can view sentiment about your business, and they can switch loyalty in a single click any time anywhere all from a mobile device. In addition, the emergence of social media sites means that customers also have a voice. They can express opinion and sentiment about products and brands on Twitter, Facebook, and review Web sites and create social networks by attracting followers, and following others.

For many CEOs, customer retention, loyalty, service and growth are top of their agenda. In addition improving operational effectiveness is also high on their priority list. The only way they can achieve this is to acquire more data. CMOs also want access to new data to enrich what they already know about customers.
New data is needed to provide insight on customer on-line behaviour for better segmentation and to understand the value of a customers' social network and not just the customer. In addition, COOs want more data to become more effective in operations.

Instrumentation is therefore being added so that operations can capture new data. With so much demand we are now in an era where data has never before been so important to business in helping to create competitive advantage.
This new 2-day seminar looks at the need to capture new data sources to add to what we already know and use Machine Learning to automatically discover, profile and catalog what is in these data sources. It then looks at how Machine Learning and Advanced Analytical techniques, such as text analyses, sentiment analysis, graph and streaming Analytics, can be used at scale on Big Data to provide new insight that helps foster growth, reduce costs and improve effectiveness for competitive advantage.

AUDIENCE

Business Analysts, Data Scientists, BI Managers, Data Warehousing Professionals, Enterprise Architects, Data Architects, CIO's, IT Managers

LEARNING OBJECTIVES

• How data and analytical characteristics can dictate the approach taken and tools needed to conduct exploratory analytics
• How to develop analytical models using supervised and unsupervised Machine Learning
• How to develop Machine Learning models at scale on Apache Spark and Hadoop
• Tools for building Machine Learning models
• Tools for deploying, monitoring and re-training Machine Learning models
• Tools and techniques for discovery, analysis and visualisation of multi-structured data
• Text and Sentiment Analysis
• Scaling text analysis to run on Hadoop and Spark
• Clickstream analysis
• Graph analysis - 4 graph analytical techniques to identify shortest path, analyse connectivity, identify communities, determine influencers and important people in social networks etc.
• Scale graph analysis on Apache Spark GraphX
• Analyse fast data in real-time using Streaming Analytics
• Deep Learning with multi-layer neural networks
• Leverage Machine Learning and Advanced Analytics quickly and easily from self-service BI reports and dashboards for access over the Web and on mobile devices

1. An Introduction to Data Exploration, Discovery and Visualisation

This session introduces the relatively area of Data Discovery and Visualisation and looks at why businesses now need.

- New data sources - Structured versus multi-structured data
- What are the different analytical workloads?
- Types of Data Science tools
- Why do businesses need this new capability? - Machine Learning Example use cases
- Skills required for Data Discovery and Visualisation
- Creating a business aligned Analytics strategy

2. Getting Started with Predictive Analytics and Machine Learning

As we move into the era of smart business, looking back in time is not enough to make good decisions. Companies have to also model the future to forecast and predict so that they can anticipate problems and act in a timely manner to compete. Predictive Analytics is a therefore a key part of any BI initiative and should be integrated into analysis, reporting and dashboards. This session introduces Predictive Analytics and how shows how it can be used in analysis and in business optimisation.

- What is Machine Learning?
- Technologies and methodologies developing Predictive Analytical models
- Using supervised learning to develop predictive models for automatic classification
- Popular predictive algorithms, e.g. Linear regression, naïve bayes, decision trees, random forest, neural networks, support vector machines
- Implementing in-Hadoop, in-memory analytics using Spark
- Data Science Workbooks using Jupyter, RStudio, Databricks Cloud and Apache Zeppelin
- Accessing data in HDFS using SQL to build models
- Accessing Spark Machine Learning algorithms from Data Mining tools
- Deploying predictive models as a service, in a container, in-analytical databases and in Hadoop

- Industrialising enterprise model deployment using MLOps platforms such as Algorithmia
- Integrating Predictive Analytics with event stream processing for automated analysis of high velocity events in every-day business operations
- Accessing Predictive Analytics from self-service BI tools and spread sheets
- Clustering data using unsupervised learning algorithms
- Speeding up model development using Machine Learning automation tools
- Data Science tools: E.g. Cloudera Data Science Workbench, IBM Watson Studio, Sagemaker
- Deep Learning: Google Tensorflow, deepsense.io, PyTorch
- Moving beyond Predictive with Reinforcement learning and RAY

3. Advanced Analytics for Multi-Structured Data

This session looks at emerging analytical technologies for multi-structured data and explores how you can use them to improve business insight. Not all analytical projects are implemented using relational database technology, especially when it comes to very large data volumes with unstructured content, semi-structured JSON or XML data, sensor data, and clickstream. This session looks at the emergence of Advanced Analytics using Big Data NoSQL Platforms like Hadoop. It looks at the approaches to analysing complex unstructured and social content and the challenges of creating valuable business insight from multiple sources of unstructured content.

- Techniques for producing insight from unstructured content
- Tools and techniques for analysing text
- Understanding the 'voice of the customer' using sentiment analytics on email and social media data
- Clickstream analysis
- Graph analysis
  o Path analytics
  o Connectivity analysis
  o Community analysis
  o Centrality analysis
  o Finding Influencers in social networks
  o Calculating follower susceptibility to be influenced

- Streaming Analytics
  - o What is data-in-motion
  - o Use cases for streaming data
  - o Time series analysis and streaming data
  - o Tools for managing streaming ingest, e.g. Apache Kafka, StreamSets, Hortonworks Data Flow
  - o Open source streaming engines –Apache Storm, Apache Spark, Apache Flink, Google Data Flow
  - o Commercial Streaming Analytics products
  - o Developing Streaming Analytics applications with no programming
  - o Modernising your architecture to accommodate streaming data
  - o Future proofing your architecture
- Artificial Intelligence
  - o IBM Watson

4. Search, BI & Big Data

This session will examine the growing role of Search in an analytical environment both as an information consumer tool for self-service BI and as a way of analysing both structured and unstructured data. Search has been incorporated into BI tools for some time, but with the emergence of Big Data as a platform for analysing unstructured information, it is taking on a major new role. Search is a simple mechanism that is familiar to most people, and opening up the interactive use of BI via Search can have enormous business benefits. Search can be used to grow the use of BI to a much wider group of users and also provide a way to extract additional insight from unstructured content.

- Why Search and BI?
- The growing importance of analysing unstructured content
- The implications of Big Data on Search and BI
- Creating Search indexes on multi-structured data
- Building dashboards and reports on top of Search engine indexed content
- Using Search to analyse multi-structured data
- The integration of Search with traditional BI platforms
- Using Search to find BI content and metrics
- Guided analysis using multi-faceted Search
- The search based analytical tools marketplace: Apache Solr (Lucene), Attivio, Cloudera Search,

Connexica, HP IDOL, IBI WebFocus Magnify, IBM Watson Explorer, Microsoft, Quid, SAP Lumira, Splunk, Thoughtspot

5. Deploying and Using Self-Service Data Discovery and Visualisation Tools

Self-service data discovery and visualisation tools are frequently sold into business departments so that local business analysts can start building their own BI reports, dashboards and applications without having to wait for IT. These tools offer the attraction of agile development and much faster time to value. When business areas buy them it often means that development starts without any IT guidance and quickly spreads to other parts of the business with little thought for integration or re-use. The result is that inconsistency and chaos can quickly set in. This session looks at best practices in deploying these tools and how to maximise business benefit through re-use and integration with Predictive and Advanced Analytics deployed in-database, in-Hadoop, in-Spark and in-Streaming Analytics platforms to leverage analytics at scale. It also looks at newly emerging OLAP on Hadoop to enable scalable multi-dimensional analysis.

- The self-service BI tools marketplace – Microsoft PowerBI, Qlik Sense, Tableau, MicroStrategy, Oracle Analytics Cloud, ThoughtSpot, Information Builders WebFOCUS, IBM Cognos Analytics, SiSense etc.
- Key features of Data Discovery and Visualisation tools
- Requirements and Best Practices for successful self-service BI
- Self-service BI tool access to Big Data via SQL on Hadoop
- Simplifying elf-service BI tool data access to multiple data stores via data virtualisation - Logical Data Warehouse
- Accessing in-database, in-Hadoop and in-Spark Predictive Analytics from self-service BI tools and spread sheets
- Accessing streaming data and real-time Analytics from self-service BI tools and spreadsheets
- Integration with Advanced Analytics in the Cloud and on-premises

INFORMATION

PARTICIPATION FEE
€ 1100

The fee includes all seminar documentation.

SEMINAR TIMETABLE

9.30 am - 1.00 pm
2.00 pm - 5.00 pm

HOW TO REGISTER

You must send the registration form with the receipt of the payment to:
info@technologytransfer.it

TECHNOLOGY TRANSFER S.r.l.
Piazza Cavour, 3 - 00193 Rome (Italy)

PAYMENT

Wire transfer to:
Technology Transfer S.r.l.
Banca: Cariparma
Agenzia 1 di Roma
IBAN Code:
IT 03 W 06230 03202 000057031348
BIC/SWIFT: CRPPIT2P546

GENERAL CONDITIONS

DISCOUNT
The participants who will register 30 days before the seminar are entitled to a 5% discount.

If a company registers 5 participants to the same seminar, it will pay only for 4.

Those who benefit of this discount are not entitled to other discounts for the same seminar.

CANCELLATION POLICY
A full refund is given for any cancellation received more than 15 days before the seminar starts. Cancellations less than 15 days prior the event are liable for 50% of the fee. Cancellations less than one week prior to the event date will be liable for the full fee.

CANCELLATION LIABILITY
In the case of cancellation of an event for any reason, Technology Transfer's liability is limited to the return of the registration fee only.

MIKE FERGUSON

MACHINE LEARNING
AND ADVANCED ANALYTICS

December 1-2, 2020

Registration fee:
€ 1100

first name ....

surname ....

job title ....

organisation ....

address ....

postcode ....

city ....

country ....

telephone ....

fax ....

e-mail ....

Stamp and signature

Send your registration form
with the receipt of the payment to:
Technology Transfer S.r.l.
Piazza Cavour, 3 - 00193 Rome (Italy)
Tel. +39-06-6832227 - Fax +39-06-6871102
info@technologytransfer.it
www.technologytransfer.it

If anyone registered is unable to attend, or in case of cancellation of the seminar, the general conditions mentioned before are applicable.

SPEAKER

Mike Ferguson is Managing Director of Intelligent Business Strategies Limited. As an analyst and consultant he specialises in Business Intelligence and Enterprise Business Integration. With over 34 years of IT experience, he has consulted for dozens of companies on Business Intelligence Strategy, technology selection, enterprise architecture, and data management. He has spoken at events all over the world and written numerous articles. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, a Chief Architect at Teradata on the Teradata DBMS and European Managing Director of Database Associates. He teaches popular master classes in Operational Business Intelligence, New Technologies in DW and BI for the Agile Enterprise, Big Data Multi-Platform Analytics, Master Data Management and Enterprise Data Governance.