



Shell Projects & Technology

FROM DATA TO KNOWLEDGE: CONVERTING DATA TO ACTIONABLE INSIGHTS

In the last couple of years, there has been an increasing focus to utilize the collected data and transform into business actions. This change is brought about by both: easier and faster acquisition of data and lower cost for storage. High computational resources serve as icing on the cake to process this ever growing data.

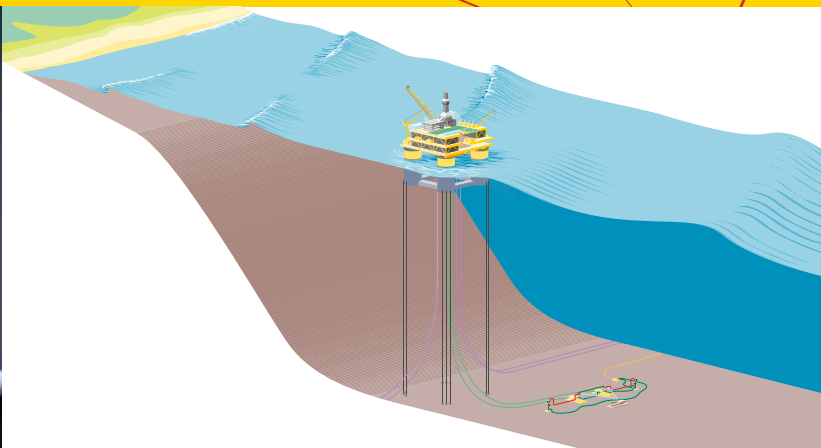
In this session we intend to discuss the best practices, challenges and opportunities for cross pollination across various domains ranging from aviation to healthcare. Additionally, we shall also hear interesting perspectives in the debate section on the value of domain knowledge for meaningful data analysis. Below is the schedule for the event.

Date: 18th December

Time: 10:00 AM - 12:00 PM

Location: Arabica, Hotel Park Plaza

Time	Topic
10:00-10:01	Session Introduction
10:01-10:25	Invited Talk 1 Speaker: Geetha Manjunath - Research Lab Manager at Xerox Research Centre India
10:25-10:50	Invited Talk 2 Speaker: Gopi Subramanian - Principal Engineer Data Science at Tyco
10:50-11:15	Invited Talk 3 Speaker: Gokul Swamy - Data Analytic Scientist at Shell
11:15 - 12:00	Debate Session
12:00-12:05	Vote of thanks
12:05-1:00	Networking Lunch



SPEAKERS

HiPC 2015, Dec 18,
10:00am–12:00pm

FROM DATA TO KNOWLEDGE: CONVERTING DATA TO ACTIONABLE INSIGHTS

Geetha Manjunath

Research Lab Manager – Xerox Research Center India

Geetha Manjunath currently heads the Data Analytics Lab at the Xerox Research Centre India (XRCI) and works on machine learning based innovative solutions in healthcare, transportation, customer care and education verticals. Before joining Xerox in April 2013, she was a Principal Research Scientist and Research Manager at Hewlett Packard Labs India for 17 years and a team leader at C-DAC Bangalore. She is a gold medallist from the Indian Institute of Science (IISc), Bangalore from where she holds a Master's Degree and PhD from Computer Science and Automation Department. She was awarded the NASCCOM IT Innovator award in 2009 and was also the winner of the 2010 Grand Challenges for Technologists, posed by MIT Technology Review under Healthcare Category. She is the author of a book on Cloud technologies titled "Moving to the Cloud" published by Elsevier Syngress Publications in Dec 2011, which is used as an Engineering college text book. She is the Chair of the Big Data Working Group and an Executive Council Member of the IEEE Cloud Computing Innovation Council for India. She holds six US patents and many more pending grant.



Actionable Insights over Healthcare Data

Abstract: Predictive Analytics for healthcare applications has immense societal benefits. It is also a very challenging task due to the high level of accuracies expected. The available data is heterogeneous and also fraught with uncertainties and noise. For example, a critical care unit are data-rich environments – physiological measurements, radiology images, lab test evaluations, clinical notes and other information are periodically recorded for each admitted patient. In this talk, I will talk about two example research projects where data analytics is being used to save lives.

Gokul Swamy

Data Analytic Scientist – Shell India

Gokul is currently a part of the computational center of expertise at the Shell India Technology Centre, Bangalore, where he is involved in developing novel data analytics algorithms for visualization, interpretation and modelling of subsurface oil and gas reservoirs. In his prior roles Gokul has been instrumental in developing advanced algorithms for image and signal interpretation of medical data. Gokul has a bachelor's degree in mechanical engineering from IIT Madras and has a PhD in electrical engineering from Michigan State University, USA. His primary research interests are in the areas of machine learning, large scale optimization and signal and image analytics. Gokul has over seven international patents and multiple journal publications in high impact scientific journals to his credit.



Detecting Reservoir Discontinuities using Sparse Binary Optimization

Abstract: Inferring reservoir characteristics from multi-phase production data and isolated measurements of reservoir properties (transmissibility, porosity, etc.) has been a hot topic of research over the past decade. This process christened assisted history matching, belongs to the class of ill-posed inverse problems wherein the parameters to be estimated far exceed the number of independent data samples. In this study we propose a sparse optimization framework, borne out of the L1 norm regularization, to delineate reservoir flow boundaries / faults which has a significant bearing on the reservoir oil production and is a key determinant for constructing reservoir flow models. The proposed methodology involves casting the problem into a bimodal parametrization and solving the ensuing combinatorial optimization problem using an evolutionary scheme coupled with a restricted boltzmann machine (RBM).

SPEAKERS

HiPC 2015, Dec 18,
10:00am–12:00pm

FROM DATA TO KNOWLEDGE: CONVERTING DATA TO ACTIONABLE INSIGHTS

Gopi Subramanian

Principal Engineer, Data Science – Tyco



Gopi Subramanian is an experienced Data Science professional with extensive experience in building machine learning applications for a variety of business verticals. In the last 15 years he has designed, developed, and lead Data Mining, Machine Learning, Text Mining, Natural Language Processing, Information Extraction and Search applications for diverse business verticals including engineering infrastructure, aviation, consumer finance, and Fire and Security. In engineering domain he has developed innovative machine learning solutions for asset optimizations and operational efficiency. In consumer finance domain, he has conceived and built innovative consumer loyalty models and has designed enterprisewide systems for personalized promotions. In text mining domain, some of his work involved analysis of short noisy engineering text arising from maintenance, repair and overhaul of engineering equipment and develop models to extract entities and conditions of parts feeding into decision support systems. He has developed search ranking schemes and associated systems for a recall based search engine for effective use within enterprises for intranet search applications. He has 16 patent applications with US and Indian Patent Office and several publications to his credit. He is the author of the book "Python Data Science Cook Book". Gopi holds an Engineering Degree in Computer Science, Master's in Computer Science and Post Graduate Diploma in Business Administration. Currently employed with Tyco as Principal Engineer Data Science. Previously he was with GE Global Research, Data Mining Lab.

Arriving at $Y = F(X)$

Abstract: Machine Learning Algorithms expect highly structured input and a well-defined problem statement. In the realms of engineering data sets structured and unstructured data are in abundance, generated by both machines and humans. Often we encounter situations where the most critical phase is the machine learning problem formulation and associated data preparation. Obtaining gold sets for training and validation is often costly and elusive. The challenge lies in effective problem formulation and associated data preparation. In this talk we will briefly talk about using free form text data for building a decision support system. During the course of the talk we will dissect the approach to formulate a machine learning problem and create associated structured data input including the gold set. Primary input to us in this problem was the free form engineering notes captured by aircraft technicians involved in Inspection, Maintenance, Repair and overhaul of Passenger Aircrafts. The extracted knowledge was fed to a decision support system which recommends an appropriate fix to a specific issue related to the aircraft.

Debate Topic: Importance of domain knowledge for data analytics

Moderator: Hima Patel (Shell)

Panelists:

Pallavi Vajinepalli (Philips)

Krishnan SH (Samsung)

Anusha Rammohan (GE)

Joy Mustafi (IBM)

Saptarshi Das (Shell)

Senthil Vadivelu (Shell)