Streaming Data Analytics
--Real time data analytics--

FORE School of Management
New Delhi

A Bold Approach To A New Business Model

Streaming data analytics is processing and analyses of data as soon as it is generated. It is analyzing data in motion rather than in batches. For most companies, whether in manufacturing or in services sector, analyzing streaming data is a compulsion to stay competitive rather than an option. For example, for supply-chain services, or taxi-services like that of Uber, the complete business model revolves around analyzing data as soon as it is generated. For Uber, most streets are market places and in each market place as demand for taxi-service goes on changing, minute to minute, resources have to be deployed/redeployed. Every online transaction must be analyzed as to its genuineness and for being non-fraudulent in the shortest possible time without causing any inconvenience to clients. In stock markets, market changes require to be tracked and settings of customer portfolios adjusted based on configured constraints, such as selling when a certain stock value is reached. On an e-commerce site, clickstream data (or logs) are to be continuously monitored; some pages may be heavily '*stormed*' by customers after marketing-campaign or during festival seasons. And, depending upon the '*crowd*', necessary computing resources (or inventory) may have to be taken care off. Real time analytics forms the backbone of many recommender engines--as soon as a customer logs in or arrives at a particular page, relevant suggestions for products must be displayed. In preventive maintenance, significant patterns of events in sensor fitted processes, like a change in pressure or temperature or change in chemical composition, need to be spotted that could indicate pending equipment or flow failure.

We make it easy to chain data pipelines, wrangle data and create fairly complex dashboards using highly reputed and high performing open source software.

1. Completely hands-on course--Students work with streaming data from business & Industry
2. Students experiment with industry standard open-source technologies—frequently used in most cloud platforms.
3. After the program concludes, we are open to a six-months free-consultancy in implementing a pilot-project on building real-time analytics system.

About the Course

In this course, students will be exposed to several Big Data open-source technologies that together constitute a stream-analytics-system. We begin with studying big data storage systems: Hadoop, several NoSQL databases and Data Lake. Emphasis is on collecting, storing and analyzing data streams. We introduce students to wrangling, processing and analyzing real time streams using Apache Spark. We also use very popular open-source technologies—Apache Kafka, KSQL (from Confluent) and Apache Flink.

Students learn to plumb data pipelines using, among others, Apache Nifi, Apache Structured Streaming and big-data storage systems. Systems with which students work include highly reputed Time series stacks such as Telegraf-Influxdb-Grafana (TIG stack) and stack to analyze logs as Elasticsearch-logstash-kibana (ELK stack). These stacks not only help in real-time analysis but also in building dashboards for real-time monitoring. The complete course is hands-on and heavily lab oriented. Almost all the technologies learnt are used in Microsoft Azure or on AWS cloud. The complete course requires the use of python. We teach python right in the beginning—students need not have any prior knowledge of working with python.

After the conclusion of the program, we are open to providing free consultancy for six-months on a pilot-project in building real time systems to analyze and monitor streaming data.

Low-code program

For those students, who wish to take a low-code approach to streaming data analytics, we have a special module. In this module, we use Graphical User Interfaces (GUIs) to learn and experiment with big-data NoSQL systems and use SQL (for example KSQL of Confluent) to analyze data-streams in and out of Apache Kafka. Students are able to plumb data-pipelines using an excellent visual framework provided by Apache Nifi. Students experiment with TIG stack and ELK stack to build complex dashboards for monitoring data-streams.

Program duration and venue

The complete program duration is 56-hours. For Low-Code program, program duration is 40 hours. Classes can be held online or in class-rooms. In online mode, classes can be held either on appointed weekdays or on Saturdays and Sundays--each class is of 2-hour duration. In class room mode full day classes are held and these can be held at FORE School of Management, New Delhi. . We can also have a mix of offline and online classes.

Program requirements

Students must have access to high-speed Internet (generally available now a days) and a lap-top with minimum of 16gb RAM. All software that we will work with are open-source and freely available. Students will also be provided with Virtual machines that have pre-configured software installed for experimentation.

Target Participants:

The program has applications in a large number of domains. Category of possible participants include Executives, engineers, decision makers, Data Scientists, Security Analysts, Business Analysts and Data Governance Teams.

Exercises and Projects

There is a heavy emphasis on exercises and projects. Students must experiment and implement systems themselves. Throughout the course students are to undertake several projects. We encourage students to use their organizational data to solve related problems.

Contacts

For any details, please feel free to contact either the Program faculty, Prof Ashok K Harnal, at 8750893093 (WhatsApp) or Prof Asif Zameer, Chair, Executive Education at 9871053303 (WhatsApp).

Program Faculty

Prof. Ashok Kumar Harnal

Ashok Kumar Harnal has worked extensively at multiple facets of Big Data Systems--Machine Learning, Deep Learning & NLP, Big-Data storage systems (Hadoop and NoSQL databases), Graph Databases, Streaming Analytics using Apache Spark, Apache Kafka, Confluent and Reinforcement Learning. He has been teaching Big Data technology since around last twelve years. Since last nine years Prof Harnal has been collaborating closely with University of California, Riverside, in a program on taking sessions on Big Data for Executives from around the World. We have trained officers from several organizations including RITES, NABARD, TechMahindra, Punjab National Bank, Central Bank of India and Union Bank of India Presently we are training officers in one another Bank. What is a matter of pride for us is that many of our students are at very high positions in Industry. My GitHub site is here. We have successfully conducted three programs on Healthcare Analytics; two programs were of three months duration and one of nine months duration. During his stay in Min of Defence, he has executed three country-wide projects on Information Systems: (a) *Raksha-Bhoomi* to computerize land records (as old as 150 years); (b) Knowledge Management of land-title related files/maps in all Defence Estates offices; and (c) Setting up of a Disaster Management organization: Archival Unit and Resource Center (AU&RC), at Delhi and Pune for safe storage of land-title related records in paper and digital forms. He has published two books (both by Tata McGraw-Hill); One on *How to program games on Computers* and the other on *Linux Administration and Applications*.

Prof Shilpi Jain

Shilpi Jain is an Area Chair and Professor of Business Administration in the Information Technology & Big Data Analytics area at FORE School of Management. She teaches graduates and executives how to align an organization's business strategy and Information Technology (IT) strategy in a dynamic business environment for achieving financial stability, sustainable growth, and operational efficiency by using scalable & secured technologies and business analytics. She has held research positions in a variety of functional areas. During her stint at Infosys Research Labs, she has designed prototypes for virtual team training platform enabling paired programming, country risk analysis framework, and dynamic resource sharing across geographies. She chairs several executive education programs on e-commerce strategies, managing business on cloud, data summarization, and business storytelling with data. She has been presenting her research in conferences of repute like

ICSE, PACIS, AMCIS, and ANZAM. Largely her current research is classified into two domains: 1) studying user behaviour on digital channels and 2) team behaviour in virtual settings. In the user behaviour context, she and her co-authors designed applied and empirical research to explore user behavior on AI Conversational Agents, Social Media channels & e-Commerce, Digital Inclusion at the Grassroots level, and adoption of disruptive technologies. Apart from that she has worked extensively with virtual teams in the organisational setting and crowdsourcing contests. At present, she is a co-chair of a funded longitudinal policy research project which aims to evaluate the impact of internet connectivity in rural India and its influence on the regularization of Rural Telecom ISPs. Prof. Shilpi's research has been published in reputed academic and case research journals, including the Journal of Business Research, Behaviour & IT, Ivey Business School, WDI Publishing, ACRC, and ACRJ, amongst others. She has presented her research at several top-tier conferences, such as ICSI, BAM, ICSE, PACIS, AMCIS, and ANZAM. Apart from conducting academic research, she concluded a policy research project titled "Digital inclusion and empowerment in Indian Handlooms In 2018. One of her teaching cases has received a 3rd place win at the WDI 25th Anniversary Case Writing Competition. And received an honorary mention in the Ivey Case Publishing Competition held in 2019.

Prof. Amarnath Mitra

Dr. Amarnath Mitra is working as an Associate Professor in the area of Information Technology and Big Data Analytics at FORE School of Management, New Delhi. Prior to joining FORE, Dr. Mitra worked as Senior Quant Analyst at BioUrja Power LLC (Texas, USA). Dr. Mitra has over five years of industry experience as an analyst and researcher with substantial exposure of working with big & high frequency data and analytics. In academics, Dr. Mitra worked as full-time faculty for over six years in management institutes such as BML Munjal University Gurugram, IMI New Delhi and IBS Hyderabad. As guest/visiting faculty he has taught in several reputed institutions like SIBM Pune, NMIMS Hyderabad, IIIT Bhubaneswar among others. Dr. Mitra has taught subjects like Data Science, Predictive Analytics, Business Analytics, Quantitative Methods, Business Research Methods, Operations Research, Econometrics, among others.

About FORE School of Management (FSM)

Foundation for Organizational Research and Education (FORE) is committed to the advancement of Management Education, Research, Training and Consultancy. Incorporated in 1981, as a non-profit institution, FORE has been working with industry and academia for developing new domains of managerial thought and education and contributing to building leaders in today's global business environment.

FORE School of Management (FSM) has been working with industry and academia for developing new domains of managerial thought and education and contribute to building leaders in today's global business environment.

Located in the heart of South Delhi, FSM provides contextual learning and helps in the development of students as thinking professionals, who have the ability to meet the future challenges of tomorrow's corporate leaders. The programmes develop multiple skills including managerial decision-making, problem-solving, analytical reasoning, communications, creativity, and innovation. FSM has been designing, developing, and conducting innovative Executive Education (EE)/ Management Development Programmes (MDPs) for working executives in India for over three decades.

FSM takes pride in its professional and high-quality faculty, modern infrastructure, technology, and resources- be it in the fields of General Management, Data Science, Human Resource, Finance, Operations, Marketing, Information Technology, Economics, and International Business.

Customized Training Program

These Programs are designed according to the specific needs of the corporate. The pedagogy used in keeping with the background, experience and aspirations of participants as specified.

Long Duration Training Program(LDPs)

Along with the above, FORE does long-duration programs like PGPM (Executive Management program), Big Data Analytics, Marketing Analytics, Healthcare Analytics. These are online or blended programs of 3 months to 11 months.

Open Training Program (OTPs)

FSM Open Training Programs (MDPs) aim to equip business managers with knowledge, skills, and attitudes for effectively responding to global developments and competitive requirements. The emphasis is on developing the ability to apply learnings efficiently and improve decision-making.