

ADVANCED TOPICS

1. Data Warehousing
2. Data Visualization
3. Natural Language Processing
4. IoT and Data Analytics
5. Deep learning and its applications using R
6. Recommender Systems

DATA WAREHOUSING

Learning Outcome Statements

- Understand the basics of data warehouse architecture
- Understand advanced data warehouse design
- Examine the issues encountered in a data warehouse project

Key Contents

Data Warehouse Architecture

- Operations System vs. DW Systems
- Data Quality
- ETL
- DW Vs. Data Mart
- Status of DW Systems

Basics of Data Warehouse Design

- ER Modelling Vs. Dimension Modelling
- Dimension Modelling – Facts and Dimensions
- Designing Star Schema
- DW Keys

OLAP using Pivot table of Excel

Advanced Data Warehouse Design

- Multiple Hierarchies
- DW BUS Matrix Architecture
- Rapidly Changing Monster Dimension
- Fact less fact tables
- Dimension Tables in Multiple Roles
- Degenerate Dimension
- Junk Dimension
- Snow flaking
- Different type of Facts

Fine Tuning Design

- Aggregates Building
- Partitioning, Indexing

Issues involved in Data Warehouse Project

- Assess Readiness
- Business Requirements
- Data Track, Analysis Track
- Deployment & Maintenance

DW Design under various scenarios (Exercises)

Advanced Data Warehouse Design Case Study I

- Designing Data Model
- Designing & Creating DW structure
- Designing & Creating Staging Area Structure
- Designing & Creating E-T-L Audit Tables
- Populating default data and static dimension
- Designing E-T-L Workflow & Loading DW
- Creating and Querying Cubes

New Trends

- Operations Data Store
- Real Time Data Warehouse
- DW from semi-structured Data
- Building Data Warehouse

DATA VISUALIZATION

Learning Outcome Statements

- Understand the concepts of cognition, perception, orientation
- Get familiar with tableau
- Learn how to create a basic dashboard
- Learn how to analyse large datasets and generate visual insights

Key Contents

- Visual Cognition
- Perception, Orientation
- Creation of Different Chart types and Interpretation
- Commonly used features in Tableau
- Principles of Analytical design
- Elimination of Visual Clutter
- How to use filters extensively
- Creation of Groups
- Create and modify Hierarchy of data for Drill ups & Drill downs
- How to create a Basic Dashboard with interactivity
- How to create a Basic Story book
- Effective display of quantitative information
- Visualization of Time Series
- How to incrementally enhance visualization using calculated fields
- Variations in Geo Spatial Analysis
- Math, Text, Logical & Date calculations
- Using Parameter controls to perform the following-
 - Dynamic Multi- Dimensions for various scenarios
 - Create What-if Scenarios
 - KPI Controls
- Quick Table Calculations
- Create special chart types
- Bollinger, Pareto, Funnel, Waterfall, Control charts, Word clouds
- Generate Visual Insights
- Separate out noise from facts.
- Rapid analysis of large data sets
- Power of Sets
- Work Sheet & Dashboard Actions
- Guided Analytics
- Formatting
- Putting it all together – In class hands on exercise: Industry example of visual
- Putting it all together – In class hands on exercise: Industry example of visual storytelling and visual analytics

NATURAL LANGUAGE PROCESSING

Learning Outcome Statements

- Gain insight on various analytical techniques used in evaluation of investment opportunities.
- Extend these concepts to the portfolio of securities and the concept of diversification, management of a portfolio.

Key Contents

Lexical Module

- Fundamental Concepts – Regular Expressions - Tokenization and Sentence
- Segmentation - Normalization - Minimum Edit Distance - Morphology
- N-Grams - Probability - Language Modeling (Evaluation & Smoothing)

Syntactic Module

- Part-of-Speech and Named Entity Tagging
- Word Classes - Rule based and Stochastic – Hidden Markov Model – Maximum Entropy
- Parsing Techniques
- Top-Down and Bottom-Up – Ambiguity

Semantic Module

- Lexical and Computational Semantics
- WordNet – Relations - Word Sense Disambiguation
- Information Retrieval
- Vector Space Model - TF-IDF – Evaluation Measures

Application Module – I

- Text Classification and Sentiment Analysis
- Naïve Bayes – PMI – Sentiment Components - Co-reference Resolution – Social Media
- Data – Tracking

Application Module – II

- Machine Translation
- Language Similarities and Differences – Lexical and Syntactic Transfer – RBMT – EBMT - SMT
- Question Answering and Summarization
- Question Classes and Processing
- Answer Extraction
- Textual Entailment
- Key-phrase Extraction
- Single and Multi-Document Summarization
- Future of NLP

IoT & DATA ANALYTICS

Learning Outcome Statements

- Understand key concepts of IoT
- IT stack needed to deal with IoT systems
- Gain an insight into the wide variety of communication infrastructure of IoT, protocols and interfaces for interacting with IoT systems
- Gain deep understanding of analytics for IoT
- Explore application of IoT in industry

Key Contents

IoT Overview

- IoT basics
- IoT reference model
- IoT reference architecture
- IoT APIs
- Analytics for API
 - Descriptive
 - Inferential
 - Exploratory
 - Predictive
 - Prescriptive

IoT Use Case Data

- Broad application of IoT across industry verticals
- IoT in smart environments

IoT Data Processing and Reference Architecture

- IoT data characteristics
- IoT data reference architecture

KPIs in IoT Analytics

- KPIs across industries
- Anomaly detection

RECOMMENDER SYSTEMS

Learning Outcome Statements

- Understand the key concepts of recommender systems
- Compute a variety of recommendations from datasets

Key Contents

- Introduction to Recommender Systems
- Types of Recommender Systems
- Related algorithms
 - Collaborative Filtering
 - Content-based
 - Hybrid models
- Evaluation of Recommender System
- Hands-on case study