



UNIVERSITY OF PISA
Department of Computer Science

Courses offered in English at the
Master Degree in
Business Informatics

Starting from Academic Year 2014/15

June 5, 2015

1 Subjects from the Business Intelligence Area

Decision Support Databases (600AA) (6 ECTS)

Semester 1

Contact Person Prof. Salvatore RUGGIERI (ruggieri@di.unipi.it)

Objectives

The course presents the main approaches to the design and implementation of decision support databases, and the characteristics of business intelligence tools and computer based information systems used to produce summary information to facilitate appropriate decision-making processes and make them more quick and objectives. Particular attention will be paid to themes such as conceptual and logical Data Warehouses design, data analysis using analytic SQL, algorithms for selecting materialized views, data warehouse systems technology (indexes, star query optimization, physical design, query rewrite methods to use materialized views). A part of the course will be dedicated to a set of case studies.

Syllabus

- Information systems and computer-based information systems in organizations.
- Decision Support System Based on Data Warehouses.
- Data Models for Data Warehouses and On-line Analytical Processing.
- Conceptual and logical design in Data Warehouses.
- Algorithms for Selecting Materialized Views.
- Data Warehouse Systems Technology: Indexes, Star Query Optimization,
- Physical Design, Query Rewrite Methods to Use Materialized Views.
- Case studies.

Data Mining (420AA) (12 ECTS)

Semester 1, 2

Contact Person Prof. Dino PEDRESCHI (pedre@di.unipi.it)

Objectives

Recent tremendous technical advances in processing power, storage capacity, and interconnectivity are creating unprecedented quantities of digital data. Data mining, the science of extracting useful knowledge from such huge data repositories, has emerged as an interdisciplinary field in computer science. Data mining techniques have been widely applied to problems in industry, science, engineering and government, and it is believed that data mining will have profound impact on our society. The course is divided into two modules. The first presents an introduction to the basic concepts of data mining and the knowledge discovery process, and associated analytical models and algorithms. The second module provides an account of advanced techniques for analysis and mining of novel forms of data, and the main application areas and prototypical case studies.

Syllabus

Module 1: Foundations

- Concepts of Data Mining and the Knowledge Discovery Process.
- Data Preprocessing and Exploratory Data Analysis.
- Frequent Patterns and Associations Rules.
- Classification: Decision Trees and Bayesian Methods.
- Cluster Analysis: Partition-based, Hierarchical and Density-based Clustering.
- Experiments with Data Mining Toolkits.

Module 2: Advanced topics and applications

- Mining Time-Series and Spatio-Temporal Data.
- Mining Sequential Data, Mining Large Graphs and Networks.
- Advanced Association, Correlation and Frequent Pattern Analysis.
- Advanced Classification, Cluster Analysis and Outlier Detection.
- Visual Analytics.
- Data Mining Languages, Standards and System Architectures.
- Social Impact of Data Mining.
- Privacy-Preserving Data Mining.
- Applications: Retail Industry, Marketing, CRM, Telecommunication Industry, Financial Data Analysis, Risk Analysis, Fraud Detection, Mobility and Transportation, Public Administration and Health.

Business Performance Analysis (417AA) (12 ECTS)

Semester 1, 2

Contact Person Prof. Salvatore RUGGIERI (ruggieri@di.unipi.it)

Objectives

The course presents techniques for Business Analytics according to two views: The process-driven view of Business Process Modeling and the data-driven view of Business Intelligence. The two views are dealt with in the two modules of the course. The first presents the main concepts and problematic issues related to the process management, where processes are understood as workflow over some basic activities, and to show some of the languages, conceptual models and tools that can help to handle the main problems in a proper way. The second module presents technologies and systems for data access, for building and analyzing data warehouses, for reporting, and for knowledge discovery in databases. The accent of the module is on the use of tools and on the analysis of application problems by means of non-trivial samples and case studies. The student will be aware and able to manage the main technologies of Business Intelligence, specifically software products for effective decision support.

Syllabus

Module 1: Business Process Modeling

- Introduction to Key Issues in Business Process Management.
- Terminology and Classification.
- Process Modeling. Conceptual Models and Levels of Abstraction.
- Rigorous Workflow Models: Petri Nets and Workflow Nets.
- Tool-supported Workflow Design and Analysis: Experimentation with Integrated Tools for Business Process Design, Analysis and Verification.

Module 2: Business Intelligence Laboratory

- Introduction: Tools for Business Intelligence.
- Data Access. Location, Format and API for Accessing Data in Text Files. Standards for RDBMS Data Connectivity.
- Extract Transform and Load. Tool for ETL. Case studies.
- Data Warehousing and OLAP. Tools for Dimensional Modeling. Case Studies.
- Tools for Reporting and Multidimensional Browsing. Case Studies
- Data Mining. Tools for Knowledge Discovery. Case Studies.

Business Intelligence and Performance Management (566AA) (6 ECTS)

Semester 1

Contact Person Prof. Nicola CIARAMELLA (ciaramella@noesis-research.com)

Objectives

The course presents a conceptual framework and a repertoire of methods for maximizing profit together with a few realistic examples. The key concept is building up a model for forecasting the market demand of company's products and a strategy for fixing sales prices and volumes. The company collects data about its customers and builds up a data repository for analyzing it, applies data mining and machine learning methods and creates a forecasting model for customers behavior, applies optimization algorithms and makes decisions about sales strategy, gets feedback from the market and repeats the cycle. A part of the course will be dedicated to a set of case studies.

Syllabus

- Decision theory:
Rational decisions, limited rationality and heuristics..
- Business Intelligence:
Client behaviors, demand forecasting, market segmentation.
- Data Mining and Machine Learning:
Applications to Business Intelligence.
- Sale price and volume optimization:
Algorithms and heuristics.
- Case study: Flight reservations,
product recommendation, web advertising optimization.

2 Subjects from the Informatics Area

Big Data Analytics (599AA) (6 ECTS) – starting in A.Y. 2015/16

Semester 1

Contact Person Dr. Fosca GIANNOTTI (fosca.giannotti@isti.cnr.it)

Objectives

In our digital society, every human activity is mediated by information technologies. Every minute, an avalanche of big data is produced by humans, consciously or not, that represents a novel, accurate digital proxy of social activities at global scale. Big data provide an unprecedented “social microscope”, a novel opportunity to understand the complexity of our societies, and a paradigm shift for the social sciences. This course is an introduction to the emergent field of big data analytics and social mining, aimed at acquiring and analyzing big data from multiple sources to the purpose of discovering the patterns and models of human behavior that explain social phenomena. The focus is on what can be learnt from big data in different domains: mobility and transportation, urban planning, demographics, economics, social relationships, opinion and sentiment, etc.; and on the analytical and mining methods that can be used. An introduction to scalable analytics is also given, using the “map-reduce” paradigm.

Syllabus

Big data sources

- Open (linked) data, Web activity data, Social network data, Social media data, Mobile phone data, Navigation GPS data, Commercial transaction data, Tourism-related data, Crowdsourcing / crowdsensing.

Big data analytics and social mining methods for:

- the discovery of individual social profiles.
- the analysis of collective behavior.
- the discovery of emotional content of text and sentiment analysis.

Big data analytics domains

- Mobility and transportation.
- Nowcasting of socio-economic indicators of progress, happiness, etc.
- Twitterology and nowcasting of social mood and trends.
- Tourism.

Ethical issues of big data analytics.

- Privacy and personal data protection.
- Privacy-preserving analytics.
- Social responsibility of data scientists.

Scalable data analytics

- Paradigms of NO-SQL databases.
- Data analysis processes with the “map-reduce” paradigm.

Database Structures and Algorithms (411AA) (6 ECTS)

Semester 2

Contact Person Prof. Giorgio GHELLI (ghelli@di.unipi.it)

Objectives

Database systems occupy a central position in our information-based society, and computer scientist and database application designers should have a good knowledge about both the theoretical and the engineering concepts that underline these systems to ensure the application performance desired. The student who completes the course successfully will be able to demonstrate advanced knowledge of the main issues related to the implementation of classical centralized relational database systems in order to be a sophisticated user of database technology and a high-performance applications developer.

Syllabus

- Architecture of a DBMS:
 - The Relational and Storage Engines.
- Permanent Memory Manager and Buffer Manager.
- Storage Structures Manager:
 - Heap and Sequential Organizations;
 - Primary and Secondary (Index) Organizations.
- Access Method Manager: The Storage Engine Interface.
- Query Manager:
 - Physical Operators for Relational Operators;
 - Physical Query Plan Generation;
 - Query Optimization.
- Transaction and Concurrency Managers.
- Physical Database Design and Tuning.

Machine Learning: Fundamentals (320AA) (6 ECTS)

Semester 1

Contact Person Dr. Alessio MICHELI (micheli@di.unipi.it)

Objectives

We introduce the principles and the critical analysis of the main paradigms for learning from data and their applications. The concepts are progressively introduced starting from simpler approaches up to the state-of-the-art models in the general conceptual framework of modern machine learning. The course focuses on the critical analysis of the characteristics for the design and use of the algorithms for learning functions from examples and for the experimental modelization and evaluation.

Syllabus

- Introduction: Computational learning tasks, prediction, generalization.
- Basic concepts and models: structure of the hypothesis space, discrete and continuous spaces, linear models, nearest neighbor, prepositional models, inductive bias.
- Neural models: Perceptron and computational properties.
Introduction to multilayer feedforward Neural Networks architectures and learning algorithms.
- Rule based models.
- Principles of learning processes and general practical aspects:
Validation, Bias-Variance analysis. Elements of Statistical Learning Theory, VC-dimension.
Ensemble learning.
- Support Vector Machines: linear case, kernel-based models.
- Bayesian and Graphical models.
- Unsupervised learning.
- Introduction to Applications.

ICT Risk Analysis (416AA) (6 ECTS)

Semester 2

Contact Person Prof. Fabrizio BAIARDI (baiardi@di.unipi.it)

Objectives

The course is divided into two parts. The first one introduces the concepts underlying the risk management of an ICT system and the analysis that have to be implemented to assess and manage this risk. The approach extends to ICT system classical risk management strategies. The second parts of the course introduces cloud systems, their enabling technologies and the economic advantages they enable. Then, the security of cloud system is evaluated to show how they change the traditional approach to increase the security of an ICT system.

Syllabus

Risk Analysis of an ICT System

- Vulnerability of a System.
- Vulnerability Analysis of ICT Systems: Analysis of Standard Components, Analysis of Non Standard Components.
- Threat Analysis.
- Attack Analysis: Complex Attacks and their Formal Description, Automated Attacks (Virus and Worms).
- Impact Analysis.
- Risk Assessment and Management.
- Countermeasures: Encryption, Intrusion Detection Systems, Firewall, Secure Programming.

Security of Cloud Systems

- Enabling Technologies.
- Architectural Models.
- Deployment Models.
- Threats of Cloud System.
- Attack against Cloud.
- Countermeasures for Cloud Systems.

Information Retrieval (289AA) (6 ECTS)

Semester 1

Contact Person Prof. Paolo FERRAGINA (ferragin@di.unipi.it)

Objectives

In this course we will study, design and analyze (theoretically and experimentally) software tools for IR-applications dealing with unstructured (raw data), structured (DB-centric) or semi-structured data (i.e. HTML, XML). We will mainly concentrate on the basic components of a modern Web search engine, by examining in detail the algorithmic solutions currently adopted to implement its main software modules. We will also discuss their performance and/or computational limitations, as well as introduce measures for evaluating their efficiency and efficacy. Finally, we will survey some algorithmic techniques which are frequently adopted in the design of IR-tools managing large datasets.

Syllabus

- Search engines.
- Crawling, Text analysis, Indexing, Ranking.
- Storage of Web pages and (hyper-)link graph.
- Results processing and visualization.
- Other data types: XML, textual DBs.
- Data processing for IR tools.
- Data streaming.
- Data sketching.
- Data compression.
- Data clustering (sketch).

Laboratory on algorithms for Big Data (588AA) (6 ECTS)

Semester 1

Contact Person Dr. Rossano VENTURINI (rossano@di.unipi.it)

Objectives

The course consists of a first part of lectures describing advanced algorithms and data structures (3 CFU), and a laboratory in the second part (3 CFU) in which the students will deploy these techniques to develop a software project. The students will select their projects among a set of proposals by major IT companies which are challenging from an algorithmic perspective. These companies will also contribute to identify/construct significant datasets that will help in testing the proposed algorithmic solutions. The course will provide the opportunity of: - facing with difficult algorithmic problems of practical interest involving big data; - evaluating the impact of efficient algorithmic solutions in the design of software managing big data; - implementing advanced software by using powerful and sophisticated libraries; - getting in touch with some companies for internships, scholarships, or thesis proposals.

Syllabus

- Compressed data structures for integers, texts, and trees.
- Randomized data structures: hashing and sketching.
- Graphs: compressed representations, navigation, and analysis.
- Geometric data structures (optional).

Peer-to-Peer Systems (261AA) (6 ECTS)

Semester 2

Contact Person Dr. Laura RICCI (ricci@di.unipi.it)

Objectives

The course presents the main methodologies and techniques for the project and the implementation of P2P systems. Both unstructured and structured overlays will be analysed. Any technique will be exemplified by a set of real applications. The course will introduce a set of tools for the simulations and the implementation of P2P systems.

Syllabus

- P2P Systems: Classification and General Characteristics.
- Unstructured P2P Overlay Networks.
- Proximity Aware Overlays: Internet Coordinate Systems.
- Cooperative Content Distribution.
- Tools for P2P Network Simulation and Implementation.

Social Network Analysis (589AA) (6 ECTS)

Semester 2

Contact Person Prof. Dino PEDRESCHI (pedre@di.unipi.it)

Objectives

Over the past decade there has been a growing public fascination with the complex “connectedness” of modern society. This connectedness is found in many contexts: in the rapid growth of the Internet and the Web, in the ease with which global communication now takes place, and in the ability of news and information as well as epidemics and financial crises to spread around the world with surprising speed and intensity. These are phenomena that involve networks and the aggregate behavior of groups of people; they are based on the links that connect us and the ways in which each of our decisions can have subtle consequences for the outcomes of everyone else. This short course is an introduction to the analysis of complex networks, with a special focus on social networks and the Web - its structure and function, and how it can be exploited to search for information. Drawing on ideas from computing and information science, applied mathematics, economics and sociology, the course describes the emerging field of study that is growing at the interface of all these areas, addressing fundamental questions about how the social, economic, and technological worlds are connected.

Syllabus

Graph theory and social networks

- Graphs.
- Social, information, biological and technological networks.
- Strong and weak ties.
- Networks in their surrounding context.

The World Wide Web

- The structure of the Web.
- Link analysis and Web search.
- Web mining and sponsored search markets.

Network dynamics

- Information cascades.
- Power laws and rich-get-richer phenomena.
- The small-world phenomenon.
- Epidemics.

Software Services (389AA) (6 ECTS)

Semester 1

Contact Person Prof. Antonio BROGI (brogi@di.unipi.it)

Objectives

The course presents the main aspects in the design and implementation of software services. After introducing the currently adopted standards for Web services, the course centers on service-oriented architectures and on the techniques for developing applications by discovering, composing and adapting existing services. The use of languages supporting the definition and the implementation of business processes via workflows is discussed. The course also describes some of the techniques employed for guaranteeing non-functional properties of services, such as quality of service and security properties. The last part of the course discusses the role of software service engineering in the more general context of service economy (and of the so-called “service science”), by illustrating the interplay between engineering and economic aspects (business models, service contracts) of services, as well as the relevance of the separation of concerns in the design of services.

Syllabus

- Protocols and Architecture of Software Services.
- Basic Standards of Web Services.
- Service-oriented Architectures.
- Introduction to the Design and Implementation of Software Services.
- Discovery, Composition and Adaptation of Software Services.
- Definition and Implementation of Business Processes via Workflows.
- Business Process Execution Languages.
- Management and Analysis of Non-functional Properties (Quality of Service, Security).
- Role of Service Engineering in the Context of the Service Economy.

Technologies for Web Marketing (537AA) (6 ECTS)

Semester 2

Contact Person Prof. Salvatore RUGGIERI (ruggieri@di.unipi.it)

Objectives

The course presents an overview of the tools provided by the Web by explaining the technologies and strategies that govern their functioning, to show how it is possible to exploit these tools to implement marketing strategies to improve the visibility, understanding the users opinion about the products and influence opinions and purchasing behaviour.

Syllabus

- Calculation of the importance of a web page by search engines and the possibility of influencing their result.
- Sorting the results shown by search engines in response to a query.
- Choice of the allocation of advertising slots on search engines and new media.
- Mechanisms that regulate online auctions and their objectives.
- Identifying individuals most influential in the social network in order to predict and influence the trend of these networks.
- Analysis of social media to determine the users opinions about products or brands.
- Customization and optimization recommendations for new products in online commerce sites.
- Strategies used on websites to attract and retain visitors.

Visual Analytics (602AA) (6 ECTS)

Semester 2

Contact Person Dr. Salvatore RINZIVILLO (rinzivillo@isti.cnr.it)

Objectives

The availability of large data sources provides new opportunities for understanding patterns and behaviors of modern society. The information from these sources requires effective visualization methods to extract meaningful information from the data and facilitate the interpretation of very complex phenomena. The objective of the course is to present an overview of basic methods and visualization techniques for effective presentation of information from different sources: structured data (relational hierarchies, trees), relational data (social networks), temporal data, spatial data and data space-time. We will present and discuss several case study scenarios with the existing methods and tools.

Syllabus

Visual Metaphors for Information

- Hierarchical and structured data.
- Relational and graph-based data.
- Temporal Data.
- Spatial data.
- Spatio-temporal data.
- Unstructured information (text).

Methods and Tools

- Overview of existing visual analytics environments.

Visual Analytics Process

- Definition of a Knowledge Discovery process.
- Framework for VA.
- Visual exploration and analytics of data.
- Case studies.

3 Subjects from the Operation Research Area

Logistics (255AA) (6 ECTS)

Semester 1

Contact Person Prof. Maria Grazia SCUTELLÀ (scut@di.unipi.it)

Objectives

The course presents the structure and functions of logistics systems, analyzing major decision problems arising in the medium/long term (tactical/strategic decisions). After an introduction to the main characteristics of logistics systems, with emphasis on distribution logistics, optimization models for decision support are discussed. Some relevant models and methods are then illustrated with the aid of appropriate software tools, and logistics case studies are presented.

Syllabus

- Introduction to Supply Chain.
- Models and Methods for Location Problems.
- Models and Methods for Transportation Problems.
- Models and Methods for the Design and Management of Distribution Centers.
- Models and Methods for Inventory Management.

Model-driven decision-making methods (601AA) (6 ECTS)

Semester 2

Contact Person Prof. Antonio FRANGIONI (frangio@di.unipi.it)

Objectives

The course will enable the student to produce and/or appropriately use software tools for the support to complex decisions (mainly at the corporate/industrial level) based on mathematical optimization techniques. The course is focussed on practical aspects of these tools. The main aim is to familiarize the students with the specific computer science aspects of these activities, such as data preparation and validation, the development of complex mathematical models, the knowledgeable use of the corresponding solution algorithms, the impact on this process of data uncertainty and the available methodologies to tackle this problem.

Syllabus

- Decision theory, decision processes.
- Architecture of decision support systems.
- Reminds to the theory of Linear Programming and Integer Linear Programming problems.
- Solvers of Linear Programming and Integer Linear Programming problems.
- Methodologies for improving performances of the algorithms.
- Data uncertainty issues within optimization methods.

Network Optimization Methods (433AA) (6 ECTS)

Semester 2

Contact Person Prof. Maria Grazia SCUTELLÀ (scut@di.unipi.it)

Objectives

The course presents the main modeling techniques and the main algorithmic methodologies for managing communication networks both at a design and at an operational level. Relevant design and operational problems for communication networks will be considered, such as QoS routing problems, location problems and resiliency problems. Then, modeling techniques and algorithmic approaches will be considered for both basic problems and NP-Hard problems.

Syllabus

- Basic network optimization problems: Models and Algorithms, Minimum Cost Flow, Multicommodity Flows.
- NP-Hard Network Optimization Problems: Models and Algorithms, Routing Models, Network Design Models. Main Heuristic Techniques, Exact Approaches.
- Applications: QoS Routing, Location Problems, Resiliency Problems.