

### <빅데이터융합학과 교과목>

번호	학수번호	교 과 목 명	시 간	학 점	비 고
1	BDC101	빅데이터자연어처리기술	2	2	
2	BDC102	텍스트마이닝	2	2	
3	BDC103	빅데이터와정보검색	2	2	
4	BDC104	추천시스템	2	2	
5	BDC106	데이터시각화	2	2	
6	BDC107	데이터분석을위한확률개론	2	2	
7	BDC108	통계적학습개론	2	2	
8	BDC113	기초인공지능	2	2	
9	BDC114	기초통계와기계학습	2	2	
10	BDC115	고급기계학습	2	2	
11	BDC116	데이터과학개론	2	2	
12	BDC117	데이터과학특론	2	2	
13	BDC118	빅데이터마이닝	2	2	
14	BDC122	빅데이터분석을위한Python	2	2	
15	BDC126	건강관리분석학의기초	2	2	
16	BDC128	빅데이터응용세미나I	2	2	
17	BDC129	빅데이터응용세미나II	2	2	
18	SWS102	정보보호이론및응용	2	2	Cross-Listing
19	SWS120	프로그래밍언어및컴파일러	2	2	
20	SWS121	보안프로그래밍기초	2	2	
21	SWS130	인공지능보안	2	2	
22	AAI101	인공지능개론	2	2	
23	AAI107	기계학습	2	2	
24	AAI110	딥러닝	2	2	
25	AAI112	자연어처리	2	2	
26	AAI114	텍스트마이닝	2	2	
27	AAI117	데이터및정보시각화	2	2	
28	AAI118	의료인공지능	2	2	
29	AAI123	인공지능 응용 세미나 I	2	2	
30	AAI122	인공지능 응용 세미나 II	2	2	

## <교과목별 교수요목>

### BDC101 빅데이터 자연어처리 기술 ( Introduction to Natural Language Processing in Big Data )

자연 언어(Natural Language)에서 추출할 수 있는 정보는 매우 풍부하다. 하지만 이러한 정보를 추출하기 위해서는 복잡한 계산과 모델링, 전문지식이 필요하며 최근 생성되는 언어의 양은 빅 데이터(Big Data) 형태로 구성되어 그 과정의 전문성을 요구한다. 본 강의는 자연어처리 기술과 이를 이용한 빅 데이터 처리 기법의 응용에 대하여 학습한다.

There is a wealth of information to be extracted from natural language, but that extraction is challenging. The volume of human language we generate constitutes a natural Big Data problem, while its complexity and nuance requires a particular expertise to model and mine. In this course, we illustrate the impressive combination of Big Data Applications and Natural Language Processing Techniques.

### BDC102 텍스트 마이닝 ( Advanced Topics in Text Mining from Big data )

텍스트 마이닝은 문자로 표현된 인간의 언어를 컴퓨터로 분석 처리하고 그 구조와 의미를 이해하는 자연어 처리 기술이다. 일반적인 데이터와는 달리 텍스트는 각 언어별로 어휘적, 문법적 독특성이 있고 그 표현의 형태가 매우 다양하고 복잡하여 일괄된 규칙으로 규정하기 힘든 경우가 많다. 또한 언어가 사용되는 환경에 따라 끊임없이 변하는 언어가 가진 복잡성 때문에 아직도 도전적 목표가 많이 남아 있는 기술 분야로 인식되고 있다. 본 강의는 빅데이터에서 텍스트 마이닝에 대한 기초적 지식 및 통계적, 규칙적 알고리즘에 대해 다룬다.

Text mining, also referred to as text data mining, roughly equivalent to text analytics, refers to the process of deriving high-quality information from text. High-quality information is typically derived through the devising of patterns and trends through means such as statistical pattern learning. Text mining usually involves the process of structuring the input text (usually parsing, along with the addition of some derived linguistic features and the removal of others, and subsequent insertion into a database), deriving patterns within the structured data, and finally evaluation and interpretation of the output. This course will cover the major techniques for text mining and analyzing text data to discover interesting patterns, extract useful knowledge, and support decision making, with an emphasis on statistical approaches that can be generally applied to arbitrary text data in any natural language with no or minimum human effort.

### BDC103 빅데이터와 정보 검색 ( Introduction to Information Retrieval in Big Data )

연구 분야 뿐만 아니라 기업에서도 빅 데이터를 다루는 기술이 다양하게 활용되고 있다. 빅데이터가 다루어지기 시작하면서 빅데이터 안에서 얼마나 데이터를 효율적으로 검색, 저장하고 안정적인 서비스를 제공하는 것이 매우 중요해졌다. 본 강의는 빅데이터에서 정보 검색 기술에 대해 다룬다.

Not only scientific institutes but also companies harvest an amazing amount of data. For the arrival of the era of big data, how to improve the retrieval time, the storage capacity and stability of system, these provide for the future construction of a guiding role. This course deals with information retrieval techniques in Big Data.

### BDC104 추천 시스템 ( Advanced Topic in Recommendation System from Big Data )

추천 시스템은 사용자의 의도 및 선호에 맞춰서 사용자의 항목을 예측하고 정보를 필터링하는 시스템이다. 추천 시스템은 최근에 매우 일반화되고 있고 다양한 용도에 적용된다. 이 강의는 빅데이터를 이용한 추천 시스템에 대해 다룬다.

Recommender systems or recommendation systems (sometimes replacing "system" with a synonym such as platform or engine) are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that a user would give to an item. Recommender systems have become extremely common in recent years, and are applied in a variety of applications. This class deals with Recommendation

System in Big Data.

#### BDC106 데이터 시각화 ( Introduction to Data Visualization in Big Data )

디지털 시대의 학습자는 영상, 디지털 매체를 비롯해 SNS, 클라우드 서비스 등을 통한 무한정의 데이터에 노출되어 있으며 이와 같은 환경의 변화에 따라 데이터 표현과 수용방식의 변화가 요구되고 있다. 다시 말해 수많은 정보들을 시각적으로 묘사하고 필요한 정보를 효율적이고 명확하게 제공하는 데이터 시각화의 필요성이 증가하고 있다. 최근 빅데이터 기술시장의 성장과 함께 시각화 기술도 중요한 기술요소로 자리 잡고 있다. 또한 전통적인 시각화 기술은 주로 시스템 로그나 실험 분석 결과 등에 대한 통계정보를 그래프로 보여주는 방식인데 반해, 빅데이터의 시각화 측면에서는 모든 데이터를 살펴보는 것에 제약이 따르기 때문에 시각화의 기술적인 요소와 더불어 데이터를 요약하고, 한 눈에 살펴볼 수 있도록 돕는 시각화 방법론적 요소의 중요성이 커지고 있다. 본 강의는 빅데이터에서 데이터 시각화 기법에 대해 다룬다.

Data visualization or data visualization is viewed by many disciplines as a modern equivalent of visual communication. It is not owned by any one field, but rather finds interpretation across many (e.g. it is viewed as a modern branch of descriptive statistics by some, but also as a grounded theory development tool by others). It involves the creation and study of the visual representation of data, meaning "information that has been abstracted in some schematic form, including attributes or variables for the units of information". This class deals with Data Visualization methods in Big Data.

#### BDC107 데이터 분석을 위한 확률 개론 ( Introduction to Probability for Data Analytics )

통계에 대한 소개와 데이터 분석을 위한 확률적 방법론에서의 역할.  
불확실성을 정량화하는 것에 대한 분석 및 시뮬레이션 기반의 방법론은 동일하게 강조된다. 시뮬레이션 메소드의 정확성을 평가하는 접근법이 논의된다. 학생들은 기본 프로그램의 사전 지식을 가지고 있어야 한다. 빅데이터 설정에서 확률과 샘플링의 적용을 논의한다.

An introduction to probability and its role in statistical methods for data analytics.  
Equal emphasis is placed on analytical and simulation-based methods for quantifying uncertainty. Approaches to assessing the accuracy of simulation methods are discussed. Students should have some prior knowledge of basic programming. Applications of probability and sampling to big-data settings are discussed.

#### BDC108 통계적 학습 개론 ( Introduction to Statistical Learning )

이 과정은 통계 학습 및 회귀 분석, 분류, 클러스터링, 차원 감소를 위한 표준 학습 기술 및 특징 추출의 원칙에 대해서 소개한다.

The course provides an introduction to the principles of statistical learning and standard learning techniques for regression, classification, clustering, dimensionality reduction, and feature extraction.

#### BDC113 기초 인공지능 ( Introduction to Artificial Intelligence )

본 강의는 인공지능에서 필수적인 기초 이론과 핵심 기술을 학습한다. 이를 위하여 기계학습, 문제 해결, 탐색 방법, 전문가 시스템, 지식 표현, 로보틱스 등 지능적 동작 시스템을 만드는데 필요한 방법과 기법들을 학습한다.

Introduction to Artificial Intelligence is emphasizing the building of agents, environments, and systems that can be considered as acting intelligently. In particular, you will learn about the methods and tools that will allow you to build complete systems that can interact intelligently with their environment by learning and reasoning about the world.

#### BDC114 기초 통계와 기계학습 ( Introduction to Statistics and Machine Learning )

기초 통계와 기계학습에서는 통계적 기계학습의 기초가 되는 중요한 아이디어와 기술들을 학습한다. 경험에 의하여 자동으로 학습하고 개선하는 시스템을 가능하게 하는 알고리즘을 만들거나 학습한다.

This major to learn the critical ideas and skills underlying statistical machine learning — the creation and

study of algorithms that enable systems to automatically learn and improve with experience.

#### BDC115 고급 기계학습 ( Advanced Topics in Machine Learning )

고급기계학습 과정에서는 기계학습 관점에서 핵심적인 요소들을 볼 것이다. 또한 기계학습 영역에서 현재 연구되고 있는 주제들과 방법들의 소개한다. 정규 네트워크, 서포트 벡터 머신, 핵심 주성분분석, 핵심 정준상관분석, 비정상 데이터 찾기 등의 학습 알고리즘들과 알고리즘들을 위한 값들이 어떻게 되는지에 관하여 논의한다. 자료들은 최근에 발표된 책을 위주로 진행한다.

The course presents the elements of kernel-based methods from a machine learning perspective. It introduces the theoretical basis for studying these methods (theory of positive definite kernels, associated reproducing kernel Hilbert spaces and techniques to construct kernel functions) and present selected topics in this area. This includes learning algorithms such as regularization networks, support vector machines, kernel principal component analysis, kernel canonical correlation analysis, anomaly detection, etc., as well as a discussion of the value of these algorithms for applications. The material is primarily based on a recent book and on research publications.

#### BDC116 데이터과학 개론 ( Introduce to Data Science )

본 과목은 학생들이 데이터 관리, 저장, 그리고 도구를 다루고 적용할 수 있도록 설계되었다. 본 과목은 데이터, 파일 기반 시스템, 관계형 데이터베이스 시스럼, 탐색과 텍스트 분석, 소셜 데이터 분석과 추론 등에 대해서 학습한다.

This course is designed to introduce students to the data management, storage and manipulation tools common in data science and will apply those tools to real scenarios. This course covers Introduction to data (data types, data movement, terminology, etc.), Storage and Concurrency Preliminaries, Files and File-based data systems, Relational Database Management Systems, Hadoop Introduction, NoSQL - MapReduce vs. Parallel RDBMS, Search and Text Analysis

#### BDC117 데이터과학 특론 ( Advanced Topics in Data Science )

새로운 도전과 데이터 분석 문제를 해결하기 위해서 데이터마이닝과 정보 분석은 연구자들에게 데이터과학으로 통합된다.

본 강의는 빅데이터를 통해 알려지지 않은 패턴을 찾기위해 예측 모델을 어떻게 개발하고 테스트하는지를 다룬다. 또한 공공 및 민간 분야의 전문가 강연 및 패널 토론을 통해 데이터 사이언스 분야에 대한 국가 정책 방향, 최신 기술 동향 및 향후 전망, 활용 성공 사례 등을 학습한다.

The Data mining and Information Analysis integrates data science to provide researcher for approaching challenging data analysis problems. Students in this course learn how to develop and test models for making predictions, to search through large collections of data for rare and unexpected patterns. Also It includes experts panel and discussion based on new computing technology, successful case.

#### BDC118 빅 데이터 마이닝 ( Big Data Mining )

정부, 사업, 연구에서 복잡한 데이터를 수집, 분석, 그리고 시각화하는 것은 중대하다. 통계학, 수학, 컴퓨터 과학에 적용된 강력한 도구들은 복잡한 데이터의 의미를 발견하는데 활용될 수 있다.

본 과목은 데이터 마이닝 원리, 분석, 데이터 시각화를 다룬다. 또한 통계학적 개념과 컴퓨터과학의 사항들을 포함한다. 학생들은 알고리즘, 컴퓨팅, 데이터 분석을 실습을 통해 다룬다.

The collection, analysis, and visualization of complex data play critical roles in research, business, and government. Powerful tools from applied statistics, mathematics, and computational science can be used to uncover the meaning behind complex data sets.

This course covers the principles of data mining, exploratory analysis and visualization of complex data sets, and predictive modeling. Following topics are statistical concepts (such as over-fitting data, and interpreting results) and computational issues. Students are exposed to algorithms, computations, and hands-on data analysis in the discussion.

#### BDC122 빅데이터 분석을 위한 Python ( Python for Big Data )

Python은 배우기 쉽고, 사용하기 쉬운 오픈소스 언어이며, 데이터 조작 및 분석을 위한 라이브러리를 가지고 있다. 본 과정에서는 Python을 통하여 빅데이터 분석에 관한 것을 학습한다.

Python is a powerful, flexible, open-source language that is easy to learn, easy to use, and has powerful libraries for data manipulation and analysis.

This course cover big data analysis using python.

#### BDC126 건강관리 분석학의 기초 ( Fundamentals of Health Care Analytics )

이 과정은 건강관리 데이터로부터 현황을 보고하는데 사용되는 분석 툴들의 다른 종류들과 사용할 수 있는 건강관리 데이터의 일반적인 유형을 자세히 살펴본다. 수업 주제는 기술적 관점(진단 코드를 캡처하고 저장하는 방법)과 분석 및 보고 관점(진단코드가 유용한 정보에서 데이터로 전환되는 방법)을 모두 고려한다. 제공되는 데이터 셋은 훈련과 숙제를 통해 더 자세하게 데이터를 탐색할 수 있는 기회를 학생들에게 제공한다.

This course looks in detail at common types of available health care data, along with different types of analytic tools used to report from that data. Class topics are considered from both a technical viewpoint (such as how diagnosis codes are captured and stored) and an analytic and reporting viewpoint (how diagnosis codes are used to turn data into useful information). Supplied data sets give students the opportunity to explore the data in more detail through in-class exercises and homework assignments.

#### BDC128 빅데이터 응용 세미나 I ( Applied topics in Big Data I )

본 강의는 교육환경에서 발생하는 로그데이터의 마이닝 기술과 데이터 마이닝을 위하여 활용되는 주요 기계학습 기법을 다룬다.

This class covers fundamental principles of data mining and machine learning for data in educational environment.

#### BDC129 빅데이터 응용 세미나 II ( Applied topics in Big Data II )

본 강의는 빅데이터 분야의 다양한 응용과 신기술에 대하여 학습한다.

This class covers various application of big data techniques and its current edge techniques.