데이터사이언스융합전공(2019학년도)

1. 학문의 정의와 교육 목표

■정의: 데이터사이언스 & 데이터사이언티스트

데이터사이언스란 공공, 산업, 민간분야에서 생산되는 데이터의 양이 폭발적으로 증가함에 따라 데이터 생태계 전반(생성•수집•분석 •활용•유통•저장)에 걸쳐 제기되는 문제를 해결하기 위한 과학적이고 체계적인 접근법을 연구하는 학문 분야이다. 좁은 의미로는 컴퓨터 도구를 효율적으로 이용하고 적절한 통계학 방법을 사용하여 데이터로부터 일반화된 지식을 추출해 의미 있는 정보를 발견하고 새로운 가치를 만들어 내기 위한 과학적인 접근법을 연구하는 학문 분야라고 할 수 있다.

미국 Syracuse 대학의 데이터사이언스 교육 프로그램의 정의에 따르면 데이터 사이언티스트(Data Scientist)란 데이터의 공유를 위해 데이터를 수집, 분석, 설계, 보존하는 전문적인 지식과 기술을 보유한 사람을 의미한다. 데이터 분석과 관리는 법/정책, 보건의료, 생명/환경, 산업/경제, 사회/복지, 방송, 문화/스포츠, 연구개발 등 사회의 모든 분야에서 필요로 하는 기술이다.

데이터사이언티스트는 문제를 통계적으로 표현하고 컴퓨터 도구를 사용하여 시각화와 데이터가공과 모형화를 구현한 후 그를 이용하여 실제적인 언어로 의미있는 결과를 만들어내는 능력 뿐 아니라 협업하는 사람들과의 소통 능력이 요구된다.

■교육 목표: 데이터 분석/활용 전문가인 “데이터 사이언티스트”의 양성

체계적이고 심층적으로 새롭게 구성한 교육과정으로 오늘날의 기술, 산업, 직업 환경 변화에 초점을 맞춰 새롭게 포지셔닝(new branding)한 데이터 전문인력 교육과정인 데이터사이언스융합전공은 학생들이 제4차 산업혁명시대를 선도할 데이터 전문가로 성장, 취업할 수 있는 기회를 제공하는 것이 주요 목표이다.

공공, 산업, 민간분야에서 생산되는 데이터의 양이 폭발적으로 증가함에 따라 데이터 생태계 전반(생성, 분석, 활용)에 걸쳐 제기되는 문제를 해결하기 위한 과학적이고 체계적인 접근법이 요구되고 있으며, 이러한 문제를 해결할 수 있는 전문가의 수요가 급증하고 있다.(Harvard Business Review (Oct. 2012) - Data Scientist: The Sexiest Job of the 21st Century / Meet the people who can coax

treasure out of messy, unstructured data (by Thomas H. Davenport and D.J. Patil)

전세계적으로 데이터(분석) 전문인력 수급은 현재 극심한 불균형 상태를 보이고 있으나, 국내 대학의 경우 데이터를 수집, 분석, 관리, 저장하는 데이터 전문가를 양성하는 학과가 드물고, 해외에서도 일부 대학에서 학사 및 석사과정을 신설해가고 있는 상황이다. 이러한 환경에서 데이터사이언스융합전공은 지난 5년간(2014-2018) 성공적으로 운영한 데이터사이언스연계전공의 경험과 지식을 바탕으로 오늘날의 기술, 산업, 직업 환경 변화에 초점을 맞춘 교육과정으로 데이터 전문인력을 양성할 비전을 가지고 있다.

글로벌 컨설팅회사인 PWC에 따르면 미국내 데이터사이언티스트를 포함한 분석능력을 갖춘 인재수요는 2015년에 이미 약 2.3백만명에 이르고 2018년에는 2.9백만으로 증가한다고 예상(www.pwc.com/us/en/library/data-science-and-analytics.html) 하고 있으며 IBM사에 따르면 데이터 전문가 수요는 2020년까지 매년 약 28% 증가한다고 예측함 (www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=IML14576USEN&) 고 있다.

급증하고 있는 데이터 전문가 수요에 부응하고 새로운 글로벌 시장 기회를 선점하는데 기여할 수 있는 교육 과정의 개설 필요성에 의해 만들어진 데이터사이언스융합전공을 통해 전 산업과 경제 주체의 의사 결정에 필요한 데이터를 수집, 분석, 시각화하여 효율적으로 활용할 능력을 갖춘 데이터 사이언티스트를 교육, 훈련시키는 것을 중요한 목표로 삼고 있다.

2. 전공과목 소개

2-1. 기초(Fundamentals)

▶ 데이터사이언스개론(Introduction to Data Science/전공핵심): 이 수업은 데이터 사이언스의 기초 개념인 데이터 처리 및 편집, 통계적인 데이터 분석, 기계학습, 데이터 시각화 및 빅데이터와 같은 주제를 컴퓨터 공학 배경이 없는 학생들이 이해할 수 있는 수준에서 소개한다. 특정 개념에 대해 심층적인 학습을 하기 보다 위의 개념에 대한 전반적이지만 실제적인 이해를 하도록 돕는다. 구체적인 사례와 실습을 통해 데이터 사이언스의 기본적인 테크닉을 습득한다.

▶ 데이터사이언스와파이썬(Python for Data Science/전공핵심/국제어수업): 데이터 분석을 위한 핵심 프로그래밍 도구인 파이썬 스크립팅의 기본 내용과 활용법을 학습한다. 특히 데이터 분석과 시각화와 관련된 파이썬의 다양한 라이브러리를

활용하는 것을 학습한다.

- ▶ 데이터사이언스와R(R for Data Science/전공핵심/전공기반/국제어수업): 본 수업은 데이터 분석의 기초가 되는 R을 소개한다. R은 데이터 처리와 그래픽 디스플레이를 포함하는 통계 컴퓨팅 및 그래픽스를 하는 언어이며, 다양한 통계적 기술을 제공한다. 이는 선행 및 비선형 모델링, 고전적 통계테스트, 그리고, 클러스터링 등을 포함한다. 본 수업은 통계 개념에 대한 이해와 과 실습을 병행하여 진행한다.
- ▶ 데이터사이언스와수학1(Mathematics Ⅰ for Data Science/전공핵심): 이 수업을 통해 학생들은 데이터 사이언스에 필요한 미적분의 핵심 개념과 이론을 습득한다. 구체적으로 미적분이론을 만드는데 필요한 현대 수학의 기본개념인 극한, 연속, 함수, 급수, 미분 및 적분의 개념을 학습한다.
- ▶ 데이터사이언스와수학2(Mathematics Ⅱ for Data Science/전공핵심): 이 수업을 통해 학생들은 선형대수의 기본개념과 응용에 대해 학습한다. 구체적으로 벡터 공간, 벡터, 선형 변환, 행렬, 연립 선형 방정식에 대해 습득한다. 학생들은 습득한 지식으로 과학이나 산업현장의 문제들을 선형방정식으로 표현하고 문제를 해결하는데 목표를 두고 있다.
- ▶ 데이터사이언스와통계1(Statistics Ⅰ for Data Science/전공핵심): 이 수업을 통해 학생들은 확률 이론의 핵심 개념을 습득하고, 이에 기반한 통계 추론을 이해한다. 다양한 확률 분포와 이의 활용에 대해 배우고 특히 베이지안 확률 및 추론을 활용한 예측 기법을 학습한다. 학생들은 이를 위해 통계 패키지와 프로그래밍 언어를 활용할 수 있게 된다.
- ▶ 데이터사이언스와통계2(Statistics Ⅱ for Data Science/전공핵심): 이 수업을 통해 학생들은 데이터사이언스를 위한 통계 분석의 개념 및 기초적인 통계기법에 대해 학습한다. 데이터의 기술, 확률분포, 표본추출, 가설검증 및 기본적인 추론통계기법의 활용을 익힌다.
- ▶ 데이터사이언스와컴퓨팅1(Computing Ⅰ for Data Science/전공핵심): 이 수업을 통해 학생들은 전산학에서 자료를 효율적으로 이용할 수 있도록 컴퓨터에 저장하는 방법인 자료구조에 대해 학습한다. 선형 구조와 비선형 구조인 스택, 큐, 덱, 그래프, 트리 등에 대해 학습한다. 학생들은 프로그래밍 언어를 이용하여 학습한 자료구조를 직접 구현해본다.
- ▶ 데이터사이언스와컴퓨팅2(Computing Ⅱ for Data Science/전공핵심): 이 수업을 통해 학생들은 어떠한 문제를 해결하기 위해 정해진 일련의 절차나 방법을 공식화한 형태인 알고리즘에 대해 학습한다. 구체적으로 알고리즘을 구현하는 방법, 알고리즘의 분류, 복잡성 에 대해 습득하고 실제문제에 기반하여 알고리즘을 디자인하고 구현하는 연습을 한다.

2-2. 응용과 실습(Applications & Practices)

2-2-1 데이터관리 트랙(Data Management)

▶ 메타데이터설계실습(Practice in Metadata Design/전공핵심/CL): 이 수업은 학생들이 구조적 전문용어사전 구축 프로세스에 대한 이해를 갖추도록 설계되었다. 학생들은 일차적으로 전문용어의 수집, 정제, 선정, 개념화에 대한 실습을 통해 메타데이터의 구조화에 대한 이해를 갖추게 된다. 이차적으로는 그룹 프로젝트를 통해 메타데이터의 속성, 속성간 관계술어, 추론규칙, RDF 트리플 구조, OWL 등에 실습을 통해 메타데이터에 관한 심화된 이해를 갖추게 된다.

▶ 차세대시스템설계론(Designing Next Generation Systems/전공핵심/CL): 이 과목을 통해 학생은 XML 기반 메타데이터 설계 및 구축, 메테이터 서비스, 웹데이터에서 메타데이터의 역할, 메타데이터 연구 등에 관한 상세한 내용을 습득할 것이다. 또한 학생은 차세대 시스템의 근간이 되는 상호운용성이 가능한 XML 기반 응용프로파일 구축 방법, 방대한 메타데이터에서 링크드 데이터의 역할도 터득하게 될 것이다. 이 과목을 마친 학생은 구문적 의미적 상호운용이 가능한 메타데이터 설계 및 응용프로파일을 구축할 능력을 갖추게 될 것이다.

▶ NoSQL데이터베이스실습(Practice in NoSQL Database Management/전공일반/국제어수업): 이 과목의 주 목적은 빅데이터를 처리하기 위해서 새로 개발된 비정형데이터베이스(NoSQL)의 전반적인 내용을 다루는 것이다. 본 과목에서 학생들은 NoSQL의 기원과 다른 데이터베이스에 비해서 어떤 특성을 지니고 있는지, NoSQL의 핵심개념, NoSQL을 개발하기 위해서 사용된 다양한 데이베이스 기술에 대해서 배울 것이다.

▶ 시맨틱웹시스템구축실습(Designing Semantics Web Systems/전공일반/CL): 이 과목의 목적은 시맨틱웹 기술을 기반으로 뛰어난 정보서비스를 제공하기 위한 시스템 구축방법론을 습득하기 위한 것이다. 구체적으로 다룰 내용은 W3C에서 제정한 Resource Description Framework(RDF), RDF 스키마, Web Ontology Language(OWL) Lite, OWL DL, OWL Full 구문구조와 시맨틱웹 시스템 구축방법론을 습득하는 것이다.

▶ 데이터베이스관리1(Database Management I/전공일반/국제어수업): 국내 및 해외 정보서비스 기관과 제공 데이터베이스분야별 종류 및 주제내용을 연구하며, 데이터베이스 관리시스템과 검색시스템 및 인터페이스 처리 등을 다룬다.

▶ 데이터베이스관리2(Database Management II/전공일반/국제어수업): Relational databases are the core of enterprise computing. Although the advent of non-relational databases has made big data processing more amenable, many established databases live in organised collections of tables in relational format. Therefore, it is essential for a data scientist to be able to extract and analyse data from relational databases. This course is designed to build on the foundations of SQL to do advanced querying and manipulation of relational data. Topics include filtering, sorting, joining, indexing and query optimisation, among others. Students will also learn to interface SQL with Python/R and appreciate its usage in applications such as data warehousing.

▶ 데이터시각화(Data Visualization/전공일반/국제어수업): 본 강의는 최근 괄목할만한 연구 및 적용이 이루어지고 있는 데이터 시각화와 관련된 제반 원칙 및 이론을 학생들에게 제시하고 경험적인 차원에서 다양한 시각화 실습 기회를 제공한다. 특히 세계 최고 수준의 시각화 애플리케이션을 활용하여 복잡, 대용량 데이터를 이용자 요구 및 의사결정 상황에 적절하게 대응하는 시각화 전략을 개발, 실행하는 구체적인 경험을 통해 실무에서의 적용성을 높이고자 한다.

2-2-2. 데이터분석 트랙(Data Analysis)

▶ 응용머신러닝(Applied Machine Learning/전공핵심/국제어수업): 본 수업은 핵심 머신러닝 과제, 알고리즘과 실제의 머신러닝 프로젝트에서 꼭 필요한 기술들을 다룬다. 학생들은 주요한 머신러닝 과제인 분류, 군집화, 회귀분석과 이러한 과제 들에서 널리 쓰이는 알고리즘에 대해 배운다. 알고리즘에 대한 기본 이론뿐만 아니라 실습을 통하여 알고리즘의 실제 응용에 대해 학습한다. 실제 머신러닝 프로젝트에서 매우 중요한 부분인 데이터 전처리, 차원 축소, 모델 평가 등에 대해서도 다룬다. 모든 실습은 파이썬 기반의 오픈 소스 머신러닝 패키지로 진행된다.

▶ 소셜데이터실습(Principles and Practice in Social Data/전공일반): 이 수업에서는 디지털마케팅을 수행함에 있어 소셜데이터 분석이 어떤 역할을 하고, 어떻게 활용 할 수 있는지 학습한다. 학생들은 "성균관대학교"에 대한 디지털마케팅을 '데이터 분석 - 전략도출 - 목표수립 - 디지털마케팅 이행 계획 수립 - 이행 - 성과측정 및 평가' 등 일련의 프로세스에 따라 수행한다. 이 과정에서 검색, 소셜버즈, SNS 등의 데이터 분석과 컨설팅 기법을 적용한 전략수립을 직접 수행하고, 체험 할 수 있다.

▶ 위치기반데이터분석(Location-based Data Analytics/전공일반): 이 수업을 통해 학생들은 지리정보시스템 및 이의 활용에 대해 학습한다. 지리정보시스템은 지리 형태에 관한 데이터를 수집, 저장하고 이를 분석, 가공하여 관련 응용 분야에 활용할 수 있는 정보를 제공하는 시스템을 말한다. 학생들은 다양한 지리 기반 데이터를 지리정보시스템 및 관련 시스템을 통해 분석하는 것을 학습한다.

2-2-3. 빅데이터 트랙(Big Data)

▶ 딥러닝개론(Introduction to Deep Learning/전공핵심/국제어수업): 이 과목은 딥러닝의 기초 지식을 배우고 실용적 기술을 습득하는 것을 목표로 하고 있으며 딥러닝의 기본 구성 요소인 신경망에 대해 소개한다. 학생들은 출력 레이어(fully connected layer), 컨볼루션 레이어 및 리커런트 레이어와 같은 서로 다른 측면을 고려한 딥신경망을 구축하는 실습을 한다. 또한 컨볼루션 신경망 (CNNs)과 리커런트 신경망(RNNs)과 같은 주요 딥러닝 기술에 대해 토론한다. TensorFlow, PyTorch, Keras와 같은 최첨단 프레임워크를 사용한

딥신경망 훈련 과제는 학생들이 딥러닝 기술을 갖춘 데이터사이언티스트가 되기 위한 준비과정이다. 이 수업을 듣는 학생들은 프로그래밍 (Python / R), 선형 대수학 및 미적분학, 그리고 머신러닝에 대한 기본 지식을 가지고 있어야 한다.

▶ 인공지능개론(Introduction to Artificial Intelligence/전공핵심/국제어수업): 이 과목은 초보자를 위한 인공지능의 기본 개념과 지능형 에이전트와 같은 현대 인공지능의 기초적인 학습을 목표로 하며 문제 해결을 탐구하고 검색 개념을 소개할 뿐만 아니라 논리 및 계획과 같은 지식 표현과 추론의 기본 사항을 탐구한다. 빠르게 성장하는 인공지능의 한 분야인 기계학습(머신러닝)에 관해 배우며 자연어 처리, 시맨틱 웹, 로봇 공학 및 컴퓨터 비전의 기초, 합리적인 에이전트와의 의사 소통, 인식 및 행동을 결정짓는 주제에 대해 공부한다. 이 과목은 입문 과정으로써 인공지능 학습을 시작하고자 하는 모든 사람들에게 적합하며 학생들은 최소한의 프로그래밍만으로 과제를 수행할 수 있다.

2-2-4. 정보학 트랙(Informatics)

▶ 정보검색론(Information Retrieval/전공핵심/CL): 정보의 축적 및 검색원리, 분류 및 색인 방법, 검색 언어 및 효율, 전산기의 이용 등을 다룬다.

▶ 의학정보학개론(Introduction to Medical Informatics/전공핵심/CL): 이 강의는 수업과제 및 활동을 완수함으로써 건강 정보 레퍼런스 기술을 익히고자 의학용어 및 관련 정보기술을 포함하여 기본적인 의학문해 문제에 대해 논의한다. 이 강의는 다양한 의료 데이터베이스 및 정보원을 찾고, 평가하고, 사용하는 방법을 이해하는 데 중점을 둔다.

▶ 의료정보시스템실습(Practice in Medical Information Systems/전공일반/국제어수업): 이 강의는 의학 및 건강 관련 데이터의 주요특징, 표준 및 관리를 비롯하여 건강정보기술의 발달로 인한 최근 이슈들, 특히, 모바일 건강 및 의료 기술, 의학정보학, 보건정보학, 의료품질향상전략, 온라인 의학정보자원 및 의학정보검색, 영상의학기술, 그리고 질병데이타 관리 및 레지스트리에 관하여 논한다. 의학 및 건강 정보 데이터베이스를 활용하기 위해, 펍메드 (PubMed)와 씨날 (CINAHL)을 실습한다.

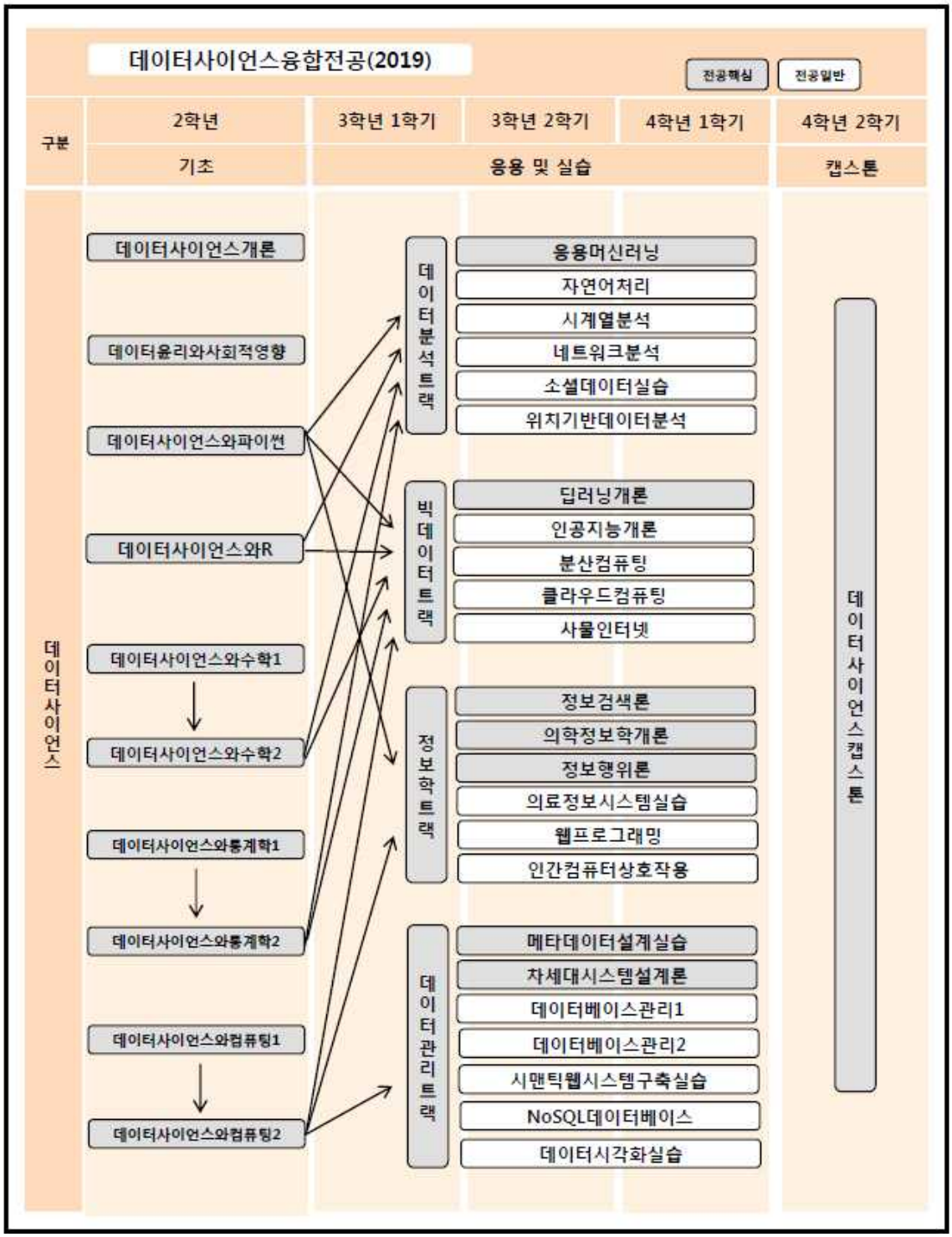
▶ 정보행위론(Information Behavior/전공핵심/CL): 인지적·심리적 측면의 정보행위를 이해하기 위한 제이론 및 통합적 접근방법을 제시하고, 정보서비스와 시스템 개발에 적용하는 기법을 검토한다.

2-3. 데이터사이언스캡스톤(Capstone Project)

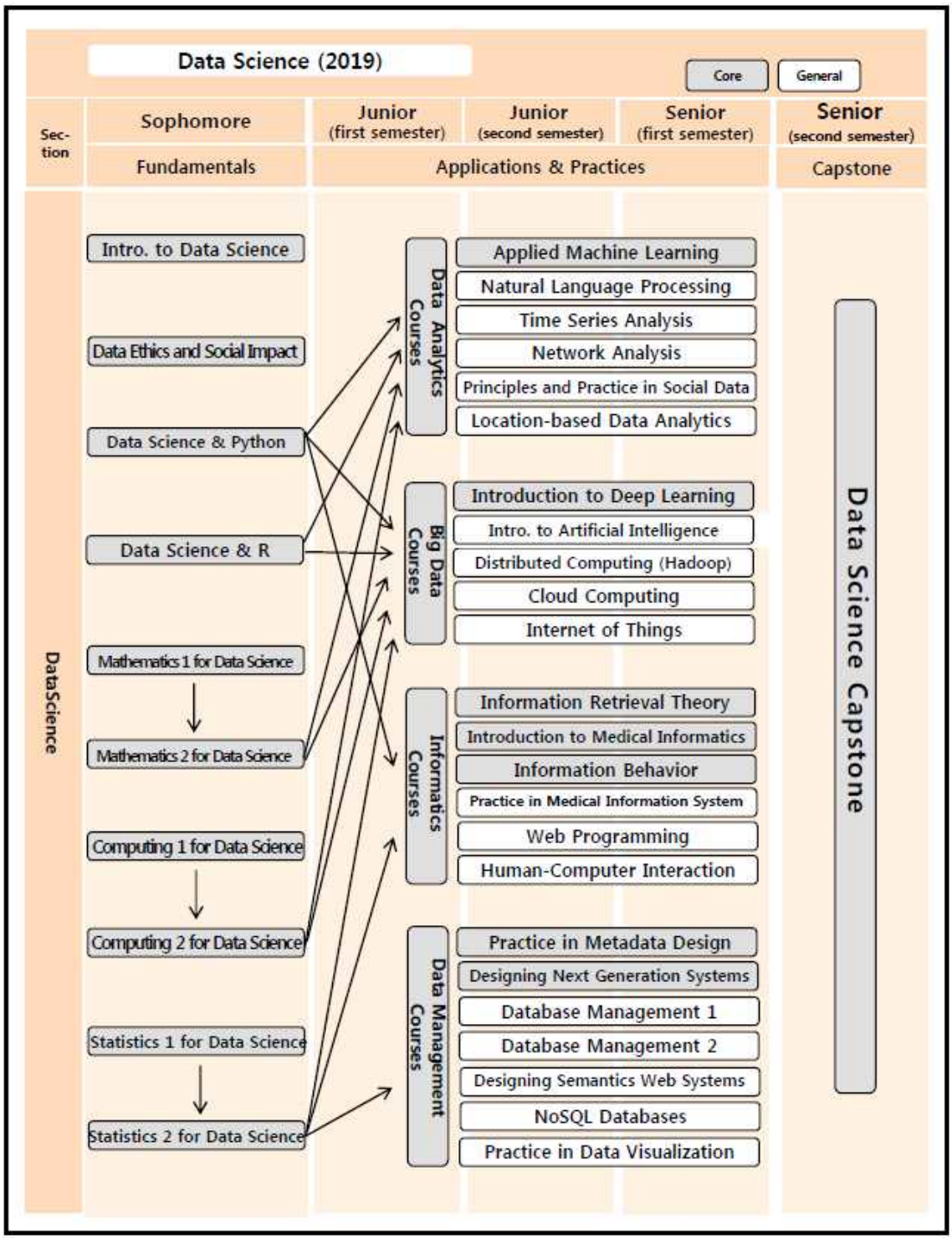
▶ 캡스톤프로젝트: 데이터 기반의 문제해결 능력을 갖춘 인재 양성과 융합적 사고 함양을 목표로 학생들이 2-3학년 때 배운 이론과 실습을 응용할 수 있는 Project-based 수업

2-4. 교과과정 로드맵

2-4-1. 데이터사이언스융합전공 로드맵(국문)



2-4-2. 데이터사이언스융합전공 로드맵(영문)



3. 비교과과정(Extracurricular Programs)

- **글로벌 데이터사이언스 캠프(Global Data Science Camp):** 하계 또는 동계 방학중 2주간 미국 시애틀에 소재한 데이터사이언스 분야의 선도적인 iSchool(Information School)인 워싱턴대학(University of Washington, Seattle)에서 오전에 데이터사이언스 수업을 듣고 오후에는 Microsoft, Amazon, Tableau, Boeing, Open Data Initiative등의 IT기업과 기관을 방문하여 데이터사이언스의 학문적 개념을 확립함으로써 학습동기를 고취시키는 동시에 데이터사이언티스트로 성장할 수 있는 기반을 제공하는 캠프 (대상은 문헌정보학과와 데이터사이언스를 전공하는 학부생이며 2019.01.20.~2019.02.02., 5차 글로벌 데이터사이언스 캠프시행)

- **부트캠프(Boot Camp):** 하계와 동계 방학 중, Excel, SPSS를 이용한 기초통계분석과 R, Python등의 컴퓨터 언어를 이용한 통계 분석 과정을 team-based, project-based로 실시하여 그 결과를 발표하고 공유함으로써 데이터를 이해하고 분석하여 원하는 결과를 도출하고 이를 시각화시켜 데이터에서 유의미한 가치를 창출하는 과정을 학습하고 팀 프로젝트를 통해 동료학생들과 의견을 교환(2019. 01. 14. ~ 01. 24 "R을 활용한 데이터사이언스" 부트캠프 시행)

- **교환학생(Student Exchange Program)**: 미국, 유럽, 중국의 명망있는 iSchool(Information School)에서 한 학기동안 공부하는 프로그램으로 취득학점이 인정되고 상대학교에서의 등록금이 면제됨(성대 등록금으로 대체). 현재 교환학생 협정이 체결된 대학;
  - ▶ University of North Carolina-Chapel Hill, U.S.A.
  - ▶ University of Pittsburgh, U.S.A.
  - ▶ University of Washington, Seattle, U.S.A.
  - ▶ University of Indiana, U.S.A.
  - ▶ Florida State University, U.S.A.
  - ▶ Penn State University, U.S.A.
  - ▶ University of Copenhagen, Denmark
  - ▶ University of Tampere, Finland
  - ▶ Peking University, China

■ **비전캠프(Vision Camp):** "비전워크숍"을 통해 문헌정보학과와 데이터사이언스 전공 진입생들이 대학 생활 전반을 계획하고 스스로 비전을 찾을 수 있도록 도와줌. 졸업한 선배들과의 "멘토-멘티 연결식"을 통해 전공과 취업과의 연계성을 공고히 하고 네트워크를 형성케 함. "선배랑후배랑" 프로그램에서는 글로벌데이터사이언스 캠프와 교환학생을 다녀온 3-4학년 선배들로부터 선발, 준비, 경험 등에 관해 듣고 묻고 대답하는 시간을 가짐. (2019년 8일, 금, 14:00, 성균관대학교 600주년기념관 6층 첨단강의실에서 시행)

4. 전공의 진로 및 전망

4.1. 관련 진로 안내

직무명	설명
① 데이터 사이언티스트	• 조직 내외부 빅데이터의 관리/활용/분석 체계를 만들고, 분석을 통해 프로세스 혁신 및 신제품 개발, 마케팅 전략 결정 등의 의사결정을 이끌어내는 직무
② 빅데이터 엔지니어	• Hadoop, NoSQL, MapReduse 등의 기술을 활용해 빅데이터 처리, LOD 구축 등 데이터 관련 개발, 기술지원, 제품 개발, 유지보수 등의 직무
③ 빅데이터 분석가	• 기업에 필요한 정형/비정형 데이터를 식별, 관리, 조작, 분석하여 의사결정에 필요한 자료를 만들어내는 직무 • 통계, 머신러닝, 텍스트마이닝 기반 데이터 분석, 분석결과 시각화 등 포함
④ 시스템 개발자	• 이용자 관점의 정보시스템, 애플리케이션 및 통합시스템 기획, 설계하는 전문가
⑤ 정보 컨설턴트	• 각종 데이터 분석을 토대로 기업이 앞으로 나아갈 방향, 해결책 등을 제시하고 관련 제품/서비스 기획
⑥ 정보기술전문가	• DB, 네트워크 시스템을 설계하여 분석하고 시스템에 맞는 소프트웨어 개발

4.2. 데이터사이언스전공 관련 자격증 안내

▶ 경영 빅데이터 분석사

빅데이터 경영 업무에 필요한 이론과 실무능력을 검증하는 자격시험 제도이다. 경영 빅데이터 분석사 시험은 빅데이터 분석 전문가를 양성하는 것이 목적이다. 빅데이터 분석 전문가는 빅데이터 이해 및 처리기술에 대한 기본지식을 바탕으로 빅데이터 분석 기획, 빅데이터 분석, 빅데이터 시각화 업무를 수행하고 이를 통해 프로세스 혁신과 마케팅 전략 결정 등의 과학적 의사결정을 지원하는 직무를 수행하는 전문가를 말한다.

- 시험주관기관: 한국경제신문 한경아카데미/(사)한국디지털정책학회
- 시험 급수: 1급(경영 빅데이터 분석사-중급)/2급(경영 빅데이터 분석사-기본)
- 시험 방법: 필기시험(1급-45문항/2급-75문항)
- 시험 과목: 1급(2과목-기획/분석), 2급(5과목-빅데이터 이해, 경영과 빅데이터 활용, 빅데이터 기획, 빅데이터 분석, 빅데이터 기술)
- 응시 자격: 1급(경영 빅데이터 분석사 2급 자격증 취득자에 한함)/2급(응시제한 없음)
- 관련사이트: http://www.bigtest.or.kr/academy/main/index.php

▶ 국가 공인 데이터 분석 준전문가

데이터 분석 준전문가는 데이터 이해에 대한 기본지식을 바탕으로 데이터 분석 기획 및 데이터 분석 등의 직무를 수행하는 실무자를 말한다.

- 시험주관기관: 한국데이터진흥원
- 시험 방법: 필기시험(50문항-객관식 40문항, 단답형 10문항)
- 시험 과목: 3과목(데이터 이해/데이터 분석 기획/데이터 분석)
- 응시 자격: 응시제한 없음
- 관련사이트: http://www.dbguide.net/da.db?cmd=snb_adsp_1

▶ 데이터아키텍처 준전문가

데이터아키텍처 직무는 기업의 모든 업무를 데이터 측면에서 처음부터 끝까지 체계화하는 것이고 데이터아키텍처 준전문가는 효과적인 데이터아키텍처 구축을 위해 데이터 요건 분석, 데이터 표준화, 데이터 모델링, 데이터베이스 설계와 이용 등의 직무를 수행하는 실무자를 말한다.

- 시험주관기관: 한국데이터진흥원
- 시험 방법: 필기시험(객관식 50문항)
- 시험 과목: 4과목(전사아키텍처 이해/데이터 요건 분석/데이터 표준화/데이터 모델링)
- 응시 자격: 응시제한 없음
- 관련사이트: http://www.dbguide.net/da.db?cmd=snb3_1

▶ 국가공인 SQL 전문가

SQL전문가는 데이터베이스와 데이터모델링에 대한 지식을 바탕으로 데이터를 조작하고 추출하는데 있어 정확하고 최적의 성능을 발휘하는 SQL을 작성할 수 있고, 이를 토대로 SQL을 내포하는 데이터베이스 프로그램이나 응용 소프트웨어의 성능을 최적하거나 이러한 성능 최적화를 지원할 수 있는 데이터베이스 객체(뷰, 인덱스 등)의 설계와 구현 등의 직무를 수행하는 전문가를 말한다.

- 시험주관기관: 한국데이터진흥원
- 시험 방법: 필기/실기시험(72문항-필기 70문항, 실기 2문항)
- 시험 과목: 3과목(데이터 모델링의 이해, SQL 기본 및 활용, SQL 고급활용 및 튜닝)
- 응시 자격: 학력/경력 기준-학사학위 이상 취득한 자/자격기준-국내외 데이터베이스 관련 자격을 취득한 자, SQL 개발자 자격을 취득한 자
- 관련사이트: http://www.dbguide.net/da.db?cmd=snb_sqlp_1

▶ 국가공인 SQL 개발자

SQL 개발자란 데이터베이스와 데이터 모델링에 대한 지식을 바탕으로 응용 소프트웨어를 개발하면서 데이터를 조작하고 추출하는데 있어 정확하고 최적의 성능을 발휘하는 SQL을 작성할 수 있는 개발자를 말한다.

- 시험주관기관: 한국데이터진흥원
- 시험 방법: 필기(50문항)
- 시험 과목: 2과목(데이터 모델링의 이해/SQL 기본 활용)
- 응시 자격: 응시제한 없음
- 관련사이트: http://www.dbguide.net/da.db?cmd=snb_sqld_1

▶ 사회조사분석사

사회조사분석사란 다양한 사회정의 수집·분석·활용을 담당하는 새로운 직종으로 기업, 정당, 지방자치단체, 중앙정부 등 각종 단체의 시장조사 및 여론조사 등에 대한 계획을 수립하고 조사를 수행하며 그 결과를 분석, 보고서를 작성하는 전문가이다.

- 시험주관기관: 한국산업인력공단
- 시험급수: 1급/2급
- 시험 방법: 1급(실기100)/2급(필기100문항60%,실기40%)
- 시험 과목: 필기 2과목(조사방법론, 사회통계)/실기 1과목(사회조사실무)
- 응시 자격: 1급(해당 종목의 2급 자격 취득 후 해당 실무 2년 종사자, 해당 실무 3년이상 종사자)/2급(응시제한 없음)
- 관련사이트: http://sa.stat.or.kr/itd/home.jsp

▶ SAS Certified Statistical Business Analyst(SCSBA)

SAS 통계 비즈니스 분석가는 대표적인 통계 분석 프로그램인 SAS/STAT® 소프트웨어를 사용하여 최적의 분석모델링을 구성하고 복잡한 통계 데이터 결과를 해석하는 능력을 소유한 분석가를 말한다.

- 시험주관기관: SAS Korea 교육센터
- 시험 방법: 필기(60문항)
- 시험 과목: 5과목(분산분석/선형회귀/로지스틱회귀/예측모형설계/모형평가)
- 응시 자격: 응시제한 없음
- 관련사이트: https://www.sas.com/en_us/certification/credentials/advanced-analytics/statistical-business-analyst.html

5. 전공 내 학회 및 스터디 소모임

동아리명	설립 목적	주요 활동 내용
문전성시	- 문헌정보학과 학회 문전성시는 전공과목 강의시간에 배우는 내용을 토대로 교과 외 활동을 통한 전공지식 함양 - 한 학기 동안 네 가지의 자유 주제 세미나와 다른 네 가지의 지정된 주제의 세미나를 통해 학과 공부와 발표 및 토론 능력 함양	- 2018 물관리 혁신 대국민 아이디어 공모전 참여 - 서울시 주관 "신공공서비스 실현 정책 제안 공모전" 장려상 수상 (상장과 상금 100만원) - 공공도서관의 현황과 발전방향에 대한 세미나 진행 세미나 주제 1. 공공도서관에서 제공하는 장애인 서비스의 문제점 및 개선방안 2. 공공도서관에서 제공하는 모바일 서비스 현황과 문제점 및 개선방안
SSR	- 문헌정보학과와 데이터사이언스 연계전공 학생들로 각종 공모전에 참여하여 기업의 문제 상황을 논리적 사고와 복합적인 의사결정 바탕의 컨설팅 연습 - 데이터 혹은 어플리케이션에 관련 솔루션 제공 능력 증진	- Pfizer 제 3회 디지털 오픈 이노베이션 공모전 참여 - 국가공인데이터분석준전문가(ADsP)와 구글 애널리틱스 자격증 준비 - 각종 기업 및 어플리케이션 개발 관련 공모전 탐색 및 준비
Bigpy	- 데이터사이언스연계전공 및 타학과 학생들과 함께 실제 데이터 분석을 하기 위해 자체적으로 Python자료 제작 및 공유 - 실제 데이터를 가지고 예시 코드 자료를 만듦으로서 머신러닝 능력 및 Python 활용 능력 연마	- 파이썬 Pandas library, matplotlib library, Scikit-learn 사용법 및 예시 코드 자료 제작 - 예시 코드 자료를 토대로 상호 피드백, 문제점 및 개선방안을 도출하여 최상의 코드 제작 목표
새싹	- 데이터사이언스연계전공학생들이 학교에서 배운 내용들을 응용하기 위해 공모전 참여 준비 - 다양한 전공의 학생들이 모여 관	- 2018 빅콘텐스트 대회 참여 '생활편의 지수만들기' 예선 진출 - 취업하고자 하는 분야의 기업 분석과 모의 면접, 예상 질문 답

	련 분야의 취업을 하고자 스터디 함으로서 다양한 분야의 기업에 대한 공부	변 준비를 통한 취업 준비

6. 데이터사이언스연계전공 선배들의 분야별 사회진출 현황(2016-2018)

데이터사이언스연계전공은 데이터와 관련된 다양한 학과의 학생들이 빅데이터 분석과 적용을 필요로 하는 사회의 많은 분야에 진출하여 활동하고 있습니다.

■ 빅데이터 분석 업무

사내 재무, 인사 등의 데이터를 분석하여 인사이트를 도출하고 경영 의사결정을 지원하며 해당 분야의 도메인 전문가인 데이터 엔지니어와 실질적으로 데이터를 핸들링하는 데이터사이언티스트와 협업하여 조직의 전반적인 의사결정을 지원하는 역할을 담당함(현대자동차 빅데이터기획팀, 삼성카드 빅데이터센터, 데이터 마케팅 코리아)

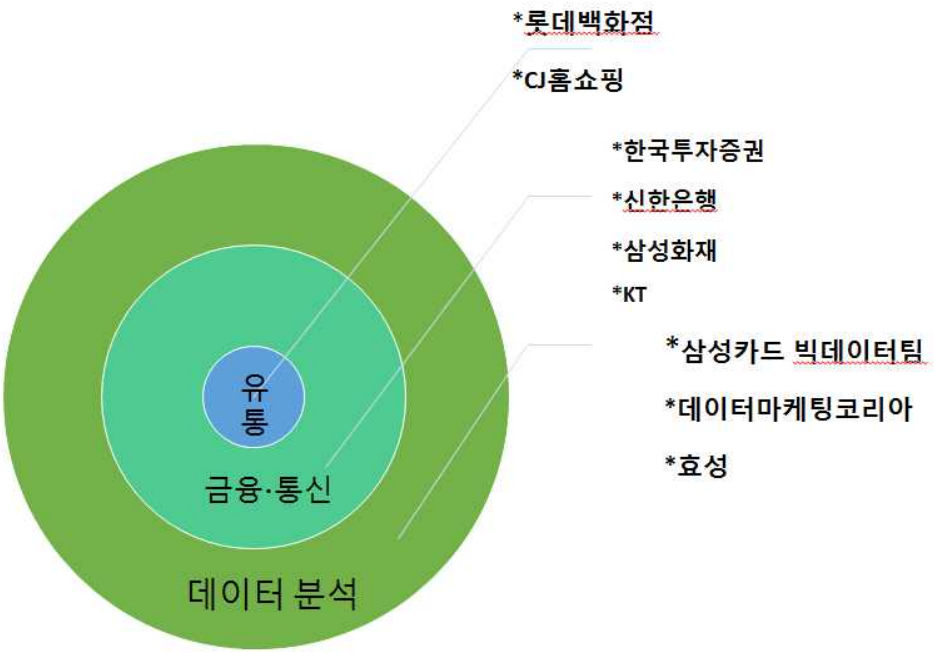
■ 금융권

시시각각 바뀌는 금융시장의 최신 데이터를 수집하고 수집된 많은 데이터들 중에서도 핵심 데이터를 분류하여 고객에게 제공하고 고객 지원과 관련된 소프트웨어를 관리하는 빅데이터 기획 마케터(한국투자증권, 우리 FIS 카드부, 신한은행)

■ 학술정보기관/데이터베이스 제공과 관리

학술정보 데이터 베이스와 같은 실제 데이터를 분석하고 온라인 DB 구축 업무와 같은 디지털 업무 지원, 데이터를 활용한 혁신적인 이용자 서비스 및 프로그램을 개발하고 학술정보 데이터베이스를 제공함(EBSCO Korea, 현대오토에버, 국립중앙도서관, LGU+)

■ 데이터사이언스연계전공 졸업생들의 취업 분야(2016-2018)



7. 데이터사이언스연계전공 재학생/졸업생 선배들의 한 마디

7-1. 전공 선배의 한 마디

- 염O필(데이터사이언스, 4학년)

저는 문헌정보학을 전공하면서, 정보학 그리고 그 정보를 이루는 데이터에 대한 관심이 많았습니다. 어떻게 하면 데이터에 접근하고 배울 수 있을까를 많이 고민하였습니다. 그러던 중 2015년 학교에 데이터사이언스연계전공이 생겼습니다. 데이터사이언스를 전공하기 전에는 단순히 툴을 잘 다루는 사람이 데이터를 잘 다룬다고 생각했습니다. 물론 툴, Python, R 등 데이터를 다루는 기술은 굉장히 중요합니다. 데이터사이언스는 이와 더불어 데이터가 무엇인지와 함께 데이터의 접근방법과 같은 기본을 탄탄하게 알려주었습니다. 이러한 기본기의 바탕 없이는 더 심도 있는 공부를 할 수 없다는 것을 깨닫게 해준 셈입니다. 가장 좋은 점은 이론적 지식을 바탕으로 툴을 함께 사용해 볼 수 있는 기회가 많다는 것입니다. 단순히 기계적으로 배우는 것이 아니라, 프로젝트를 통해서 우리가 직접 실습한 것이 가장 기억에 남습니다. 이런 프로젝트 진행과정동안 교수님들은 꼼꼼히 피드백을 주셨고 다른 학우들과는 서로 의견을 공유하고 이견을 조율했습니다. 데이터사이언스는 현재 다양한 전공을 하는 학우들이 함께합니다. 이 때문에 다른 관점을 배우며 서로서로 윈윈하고 있습니다. 또한, 데이터사이언스만의 교환학생 프로그램도 매우 잘 갖춰져 있기 때문에 데이터사이언스가 먼저 시작된 미국과 유럽의 좋은 대학에서 공부하며 글로벌마인드를 형성할 수 있다는 장점이 있습니다. 데이터를 다뤄보고 싶고, 직접 경험해보고자 한다면 데이터사이언스를 전공하는 것이 좋은 자산이 될 것입니다. 어떤 전공을 하더라도 이 전공 하나를 통해서 엄청나게 많은 것이 달라지지는 않습니다.

하지만 이를 통해서 새로운 길을 모색할 가능성, 크게 변화할 가능성을 찾았습니다. 데이터는 앞으로 살아가며 무엇을 하든 놓쳐서는 안 될 분야라고 생각합니다. 지금, 데이터를 조금이라도 아는 것은 앞으로 우리들의 삶에 많은 도움이 될 것입니다.

7-2. 데이터사이언스연계전공 졸업한 동문 선배의 한 마디

- **박O한(데이터사이언스연계전공/원전공 문헌정보학과 11학번)**

안녕하세요. 저는 학부시절 문헌정보학과와 데이터사이언스연계전공을 공부했고, 현재 현대자동차그룹 의왕연구소에서 인프라 업무를 하고 있습니다. 지금 하는 일에도 도움이 되지만, IT계열로의 취업을 준비하던 때 특히 학부시절 문헌정보학과와 데이터사이언스연계전공에서 수강한 과목들과 경험이 큰 도움을 주었습니다. 요즘은 어떤 업종이든지 간에 데이터의 중요성이 상당히 큽니다. 그 데이터를 마음대로 정제하고 재조합하고 시각화하여 인사이트 도출까지 끌어내본 경험은 면접 때마다 '데이터사이언스' 라는 다소 생소한 전공에 쏟아지는 질문에 답하기에 충분했습니다. 저의 일은 현대자동차그룹의 데이터센터, 생산사업장, 국내주요사옥, 해외법인 등의 네트웍 인프라를 구축, 관리, 운영 하는 일입니다. 이미 아시는 것처럼, 이제는 인프라의 영역도 더 이상 하드웨어의 그림자 안에 있지 않습니다. 아마존의 AWS나 마이크로소프트의 AZURE 같은 거대 도커 플랫폼(Docker Platform)이나 클라우드에 힘입어 소프트웨어를 통해 통제하는 인프라의 비중이 커지고 있습니다. 학부시절의 경험 덕에, 저는 기존의 인프라 운영방식에 자동화와 효율화를 더할 수 있는 무기를 가질 수 있게 되었습니다. 취업을 하든, 창업을 하든, 프리랜서를 하든지 간에, 이제 데이터를 갖고 노는 기술, 그리고 기초적인 컴퓨터 기술은 선택이 아니라 필수인 시대가 되었습니다. 성급하게 여러분의 진로를 선택하는 어리석은 짓은 하지 마시길 바랍니다. 대신에 안하면 손해인 좋은 기회를 놓쳐서도 안된다는 말씀을 드리고 싶습니다.

- **홍O택(데이터사이언스연계전공/원전공 중어중문학과 11학번)**

중어중문과 원전공에 데이터사이언스를 복수전공했습니다. 현재 KT 안양마케팅부에서 마케팅매니저로 근무하고 있습니다. 마케팅 매니저는 상권분석 및 담당 대리점의 수익성과 여러 지표를 관리하는 직책입니다. 대리점 관리를 잘하기 위해서는 다양한 데이터를 볼 줄 아는 능력을 길러야합니다. 또한 다양한 사람들과 만나서 협업해야 하기 때문에 열린 소통능력 또한 중요한 직업입니다. 저는 데이터사이언스를 전공하면서 여러 차례 팀 프로젝트를 경험할 수 있었습니다. 또한 팀프로젝트를 통해서 다양한 데이터를 다루는 경험을 했으며 이는 현재 직무에 많은 도움이 되었습니다. '빅데이터"의 개념이 사회에 등장한지 꽤 오래된 것 같지만 아직 현장에서는 넘치는 데이터를 제대로 활용하지 못하고 있다는 생각이 듭니다. 저 또한 많이 부족해서 지속적으로 공부해야 되는 부분입니다. 제가 복수전공을 시작했을 때보다 데이터사이언스연계전공의 커리큘럼이 점점 더 탄탄해지고 있다고 생각합니다. 이렇게 다양한 데이터를 다루는 전공은 우리나라에서 드물다는 자부심을 가지고 여러 수업에서 한 단계씩 성장하는 발판을 마련하시기 바랍니다.

8. 2019학년도 "데이터사이언스융합복수전공 신청"

■ **데이터사이언스융합복수전공 신청 안내**

▶ 복수전공 신청 기간(2019.01.15.-01.21)에 GLS로 신청(인원: 기본적으로 무제한이지만 신청자가 너무 많은 경우 학점으로 제한)

▶ 의대, 약대, 반도체시스템공학과 제외

■ **데이터사이언스연계전공에서 데이터사이언스융합복수전공으로 전환**

▶ 복수전공 신청 기간(2019.01.15.-01.21) 중 GLS에서 데이터사이언스연계전공 취소 후 데이터사이언스융합복수전공 신청(인원: 기본적으로 무제한이지만 신청자가 너무 많은 경우 학점으로 제한) ※**연계전공에서 융합전공으로 자동 전환되지 않습니다.**

■ **데이터사이언스융합복수전공(일반형)에서 데이터사이언스융합복수전공(심화형)으로 전환**

▶ 데이터사이언스융합복수전공(일반형/총 42학점 취득) 전공자는 한학기 지나고 복수전공 신청 기간 중 GLS에서 데이터사이언스융합복수전공(심화형/총 52학점 취득)을 신청할 수 있습니다.

■ 2015년 1월~2018년 7월에 데이터사이언스연계전공을 신청한 학생들이 데이터사이언스융합전공으로 전환하지 않는 경우 데이터사이언스연계전공 졸업학점(핵심12/일반24)으로 데이터사이언스연계전공 학사학위를 취득할 수 있습니다. 연계전공 학생들도 2019학년부터는 융합전공 교과과정으로 공부하게 됩니다.

※데이터사이언스융합전공 교과과정은 학점취득이 <u>절대평가</u> 방식입니다.

9. 2019학년도 "데이터사이언스융합전공" 자체개설과목 목록

(신):2019 신규과목

학기	자체개설과목		계
1	핵심	일반	9
	인공지능개론(신) 데이터사이언스개론(DSC2001) 데이터사이언스와파이썬(DSC2004) 데이터사이언스와수학1(신) 데이터사이언스와통계학1(신) 데이터사이언스와컴퓨팅I(신)	데이터베이스관리1신) 위치기반데이터분석(신) 소셜데이터실습(신)	
	자체개설과목		
2	핵심	일반	10
	딥러닝개론(신) 응용머신러닝(신) 데이터사이언스개론(DSC2001) 데이터사이언스와R(DSC2005) 데이터사이언스와수학2(신) 데이터사이언스와통계학2(신) 데이터사이언스와컴퓨팅2(신)	데이터베이스관리2(신) 데이터시각화실습(DSC3002) 의료정보시스템실습(DSC3008)	
계	13	6	19

■ 졸업요건

항목		요건		
		전공핵심	전공일반	소계
전공학점	단일전공	24	36	60
	일반형 복수전공	24	18	42
	심화형 복수전공	24	27	51
수료학점	(120) 학점			

※ 데이터사이언스융합전공 수료증(Certificate) 취득학점: 21학점(자체개설과목 12학점이상+나머지 CL 학점)

10. 더 궁금하면 직접 문의해보세요!

전공에 대한 궁금한 사항들을 전공사무실, 전공 홈페이지, 전공상담(면대면 또는 e-mail)을 통하여 해결

▶ 데이터사이언스 전공사무실: 760-1175
다산경제관 2층 32202호

▶ 전공 홈페이지 : http://cihe.skku.edu/feature/ischool01.do (국문)
http://cihe.skku.edu/en/feature/ischool01.do (영문)

11. 데이터사이언스융합전공 교수/ 2019 담당과목

교수 성명	전공분야	2019 담당과목	연구실 위치 및 e-mail주소
고영만 교수님	정보학	• 메타데이터설계실습 (C/L)	호암관 7 층 50701호 ymko@skku.ac.kr
오삼균 교수님	데이터베이스(웹)	• 데이터베이스관리1 • 시맨틱웹시스템 구축실습 (C/L)	교수회관 3 층 40314호 samoh21@gmail.com
심원식 교수님	정보서비스 (학술정보)	• 데이터사이언스개론 • 데이터시각화실습	교수회관 2 층 40226호 wonsik.shim@gmail.com
이용정 교수님	정보서비스 (건강정보)	• 의료정보시스템실습 • 의학정보학개론(C/L)	교수회관 5 층 40514호 yjy4617@gmail.com
주영준 교수님	데이터마이닝/ 인포메트릭스	• 데이터사이언스와 파이썬 • 응용머신러닝	교수회관 4층 40419호 yzhu@skku.edu
Gaya Nadarajan교수님	Informatics/ Semantics	• 인공지능개론 • 데이터 사이언스와R • 데이터베이스 관리2 • 딥러닝개론	
이진형 겸임교수님		• 소셜데이터실습	
최근호 겸임교수님		• 위치기반데이터분석	
엄태관 초빙교수님		• 데이터사이언스와수학1 • 데이터사이언스와수학2 • 데이터사이언스와통계1 • 데이터사이언스와통계2	
안용학 초빙교수님		• 데이터사이언스와 컴퓨팅1 • 데이터사이언스와 컴퓨팅2	