



数据科学与商业分析

Data Science and Business Analytics



#1. 课程背景及简介



数据科学与商业分析是一个多领域交叉学科，以商业知识为基础，通过数学统计、编程等手段，从数据分析出发，对行业、企业情况建立有效分析模型，深度挖掘数据价值，得出结论最终作为商业决策的科学依据。因此，全面专业的财经知识和扎实的数据分析能力缺一不可。

每天都会产生数百万亿字节的数据，商业世界见证了数据的快速变化，从结构化内部数据(如销售数据)和结构化外部数据(调查)，到非结构化内部数据(电子邮件)和非结构化外部数据(产品评论)。随着越来越多知名企业开始依赖数据做出关键决策，对数据分析人才的需求与日俱增。第三方权威统计，2017年数据分析招聘需求提高 150%，平均薪资增幅高达 10%。因此，蓬勃发展和令人垂涎的数据科学和商业分析就业市场需要学生掌握最前沿的商业分析知识。

在本课题的学习中，教授将会带领大家共同探讨有关商业数据调查和分析的理论及应用知识。学生有机会利用所学的相关知识分析现存的商业问题。在课程中，学生还将了解数据科学研究方法论的应用，教授会通过真实的公司商业分析数据，引导大家探究风险管理的预测性模型、市场模型、分段模型、预测和仿真模型等模型的分析及应用。学生可以在媒体娱乐领域中选择各种有趣的 PBL 项目议题，如电影票房预测，名人社交网络，或在其它相关内容。

#2. 学习目标



本课程将解决许多挑战，如：

- ★ 学习如何快速有效地进行商业数据分析并解决实际商业问题
- ★ 分析现存的实际商业问题并运用相应模型解决这些问题
- ★ 对相关领域的基础性理论知识和应用有系统的认知
- ★ 学习商业数据算法并探讨如何促进企业未来发展的方法

#3. 任课教师信息



Prof. N Z F

目前担任弗吉尼亚大学麦金塔尔商学院终身教授，研究核心是娱乐营销，数字媒体，统计和经济计量方法，机器学习和业务分析。她利用计量经济学、统计学和机器学习的方法分析大数据。教授曾在《市场研究杂志》和《市场科学杂志》发表多篇专业文章，荣获 2018 终身出版电影产业研究学术贡献奖及微星研究奖、英联邦优秀教师奖提名、弗吉尼亚大学教学奖和管理科学年度最佳评审奖等多个奖项。

4. 课程设置



周期	时间	课程设置内容	课时
第一周 学习指南 教授及助教辅导	1 月 28 日 周六	什么是 PBL 教学方法	1
		PBL 教学的常见形式	1
	1 月 29 日 周日	教授课-1 交叉学科PBL 课程设计及知识点学习 学习目标：基于项目式学习的主要步骤：确定重要和有趣的主题，了解项目大环境，确定研究目标，为最终项目引入数据源 描述：基于问题的学习 (Problem-Based Learning, PBL)：是一种强调动手实践和主动学习，以学生为中心，专注于研究和解决现实世界中复杂问题的教学方法。PBL 的几个主要特点是，学习是由具有挑战性的开放性问题来驱动，没有一个“正确”答案；研究问题或案例有特定背景；学生自我指导，积极调研，努力解决问题；学习过程中，学生需要确定一个关键问题，商定实施一个问题解决方案；教师扮演学习促进者角色，指导学生学习过程，促进探究进程。 Problem-Based Learning (PBL)的9个步骤：一般而言，PBL 包括 9个步骤（在实际应用过程中，可根据问题、学生等情况做相应调整）：（1）教师为学生将要做的项目提供真实参考并设定好步骤；（2）学生担任项目设计者的角色，可能做一个展示或比赛；（3）学生们讨论并收集他们项目所需要的背景信息；（4）教师和学生共同协商评估项目的标准；（5）学生收集项目所需要的材	3



第二周 教授及助教辅导		料；（6）学生创建他们的项目； （7）学生准备展示项目；（8） 学生讲演其项目；（9）学生根据 步骤（4）中建立的标准反思过程 并评估项目。	
	1月30日 周一	助教课-1 知识点查漏补缺	2
	1月31日 周二	教授课-2 制定个人项目方向 学习目标：数据主要类型；大数据的特点；数据的来源；收集数据的方法；数据分析管道；数据预处理 描述：数据分析指用适当的统计、分析方法对收集来的大量数据进行分析，将它们加以汇总和理解并消化，以求最大化地开发数据的功能，发挥数据的作用。数据分析是为了提取有用信息和形成结论而对数据加以详细研究和概括总结的过程。	3
	2月1日 周三	助教课-2 知识点查漏补缺	2
	2月2日 周四	教授课-3 交叉学科课程知识点学习 学习目标：数据方法拟合；分析方法的类型；评估和比较模型性能 描述：数据分析的目的是把隐藏在一大批看来杂乱无章的数据中的信息集中地提炼出来，从而找出所研究对象的内在规律。在实际应用中，数据分析可帮助人们做出判断，以便采取适当行动。数据分析是有组织有目的地收集数据、分析数据，使之成为信息的过程。这一过程是质量管理体系的支持过程。在产品的整个寿命周期，包括从市场调研到售后服务和最终处置的各个过程都需要适当运用数据分析过程，以提升有效性。例如设计人员在开始	3



		一个新的设计以前，要通过广泛的设计调查，分析所得数据以判定设计方向，因此数据分析在工业设计中具有极其重要的地位。	
	2月3日 周五	助教课-3 知识点查漏补缺& 跟进个人项目调研进度	2
	2月4日 周六	教授课-4 互动与项目设计跟进答疑	1.5
	2月6日 周一	助教课-4 跟进个人项目调研进度	2
	2月7日 周二	教授课-5 交叉学科课程知识点学习 学习目标：网络分析和文本分析 描述：沃尔玛经典营销案例：啤酒与尿布 “啤酒与尿布”的故事产生于 20 世纪 90 年代的美国沃尔玛超市中，沃尔玛的超市管理人员分析销售数据时发现了一个令人难于理解的现象：在某些特定的情况下，“啤酒”与“尿布”两件看上去毫无关系的商品会经常出现在同一个购物篮中，这种独特的销售现象引起了管理人员的注意，经过后续调查发现，这种现象出现在年轻的父亲身上。1993 年美国学者 Agrawal 提出通过分析购物篮中的商品集合，从而找出商品之间关联关系的关联算法，并根据商品之间的关系，找出客户的购买行为。艾格拉沃从数学及计算机算法角度提出了商品关联关系的计算方法——Aprior 算法。沃尔玛从上个世纪 90 年代尝试将 Aprior 算法引入到 POS 机数据分析中，并获得了成功。	2
	2月8日 周三	助教课-5 跟进个人项目调研进度	2
	2月9日 周四	教授课-6 交叉学科课程知识点学习 学习目标：有趣的发现；推导管	2

第三周 教授及助教辅导 未来展望		理影响；撰写可发表的项目报告描述：每当我们在讨论行业发展的商业趋势，不得不考虑行业商业数据、行业规模、发展速度、盈利空间、成本和利益等等相关维度；即使做简单的产品（运营）竞品分析时候也不得不收集竞品行业数据来做报告，脱离了数据来判断我们的决策没有任何依据，完全凭借自我感觉。	
	2月10日 周五	助教课-6 知识点查漏补缺&指导个人项目成果展示	2
	2月11日 周六	教授课-7 教授点评个人项目成果	1.5
	2月12日 周日	升学与就业方向展望 个人规划及发展建议	1 1
总课时	32		

#5. 阅读材料



- ★ Business Analytics at a Glance
- ★ Data Analytics: Practical Guide to Leveraging the Power of Algorithms, Data Science, Data Mining, Statistics, Big Data, and Predictive Analysis to Improve Business, Work, and Life
- ★ HBR Guide to Data Analytics Basics for Managers

#6. 项目主题



本课程使用 PBL 教学法，PBL 即项目式学习，是一种以学生为中心的教学方法，教师提供关键素材构建学习环境，学生通过在此环境里解决一个开放式项目的经历来学习。以下为本课程可选的项目主题：

- 如何准确预测电影票房？
- 如何为一部电影挑选男女演员？
- 如何设计更好的电影预告片？



英文版教学大纲



Course Title	Data Science and Business Analytics
Credit Hours	32 (one credit hour is 45 minutes)
Course Objectives	<p>This class will address many challenges such as:</p> <ol style="list-style-type: none">1.Learn how to quickly and effectively analyze business data and solve practical business problems2.Analyze existing practical business problems and use corresponding models to solve these problems3.Systematic understanding of basic theoretical knowledge and application in related fields4.Learn business data algorithms and discuss how to promote the future development of the enterprise
Course Description	<p>This course will offer students a hands-on experience of project-based learning (PBL), where students will learn by actively engaging in real-world and personally meaningful projects related to Data Science and Business Analytics.</p> <p>This course will lead students throughout every step in PBL, from selecting project topics, data acquisition, data analysis, to analytic methods, result interpretation, and project report writing. Students may select among various engaging topics on media entertainment analytics, such as movie box office forecasting, celebrity social networks, or design any topics of their own interests in this or many other domains. In the end, each student should be able to develop a solid understanding of PBL and derive an interesting project report publishable in a journal or featured on a resume.</p>



Course Background

This course will offer students a hands-on experience of project-based learning (PBL), where students will learn by actively engaging in real-world and personally meaningful projects related to Data Science and Business Analytics.

With multi-quintillion bytes of data generated each day, the business world has seen data rapidly shifting from structured internal data (such as sales data) and structured external data (surveys), to unstructured internal data (emails) and unstructured external data (product reviews). Thus, the booming and coveted data science and business analytics job market demand students to hold cutting-edge knowledge on business analytics.

This course will lead students throughout every step in PBL, from selecting project topics, data acquisition, data analysis, to analytic methods, result interpretation, and project report writing. Students may select among various engaging topics on media entertainment analytics, such as movie box office forecasting, celebrity social networks, or design any topics of their own interests in this or many other domains. In the end, each student should be able to develop a solid understanding of PBL and derive an interesting project report publishable in a journal or featured on a resume.

	Topics
Module 1	<u>Project-based Learning (PBL) and Topic Selection</u> Project-based learning and major steps Identifying important and interesting topics Understanding the literature Determining research objectives Introducing data source for final project
Module 2	<u>Data Acquisition and Pre-processing</u> Major type of big data Characteristics of big data Where to find data How to collect data Data analysis pipeline Data pre-processing
Module 3	<u>Methodology and Data Analysis</u> Data-method fit Types of analytics methods Evaluating and comparing model performance
Module 4	<u>Examples of Analytic Methods for the Project</u> Network analytics Text analytics

Module 5	Findings and Report Writing
	Discovering interesting findings
	Deriving managerial implications
	Writing publishable project report

Required Readings
Business Analytics at a Glance.

- Suggested list of the topics for the final project**
1. How to accurately forecast movies' box office?
 2. How to select actors and actresses for a movie?
 3. How to design better movie trailers?

Class Expectation
Business Analytics is a wonderful career, and continues to grow strongly. More and more companies are using data analytics to understand and drive their businesses. If you become good with these skills you will always be sought after as an employee.