Designing a Healthcare Data Analytics Course: A Contextual Active Learning Approach

Completed Research

Rachida F. Parks
Quinnipiac University
rachida.parks@quinnipac.edu

Fay Cobb Payton
North Carolina State University
fcpayton@ncsu.edu

Abstract

Despite calls for next-generation business analytics professionals, there is a lack of curricular modules that emphasize context-specific learning of analytics pertaining to healthcare. Healthcare analytics can provide different stakeholders the ability to use their data to improve patient care and engagement, evidenced-based interventions, financial efficiency and operational effectiveness. Such course has not been widely offered in business schools, hence there is no established curricular material or pedagogical research that presents a strategized curriculum. To address this need, this article discusses the use a pedagogical approach that emphasizes active and contextual learning in module design. We develop a set of new curricular modules within an interdisciplinary course - that integrates information systems and medical informatics - using a contextual active learning approach. We discuss the findings and challenges, and conclude with implications for future research.

Introduction

As the healthcare system moves from fee-for-service to fee-for-value, it became clear that much effort is needed to align quality and outcomes with cost and utilization (Sulecki et al. 2016). Such business model shifts are forcing stakeholders — including providers, payers, researchers, and governmental organizations — to consider new and innovative ways to deliver care, increase efficiencies, and improve clinical outcomes. Hence, data analytics is becoming a crucial and required operation for healthcare organizations.

With healthcare organizations becoming more data driven, health associations, such as the American Health Information Management Association (AHIMA), are calling for professionals to effectively analyze data, interpret insights and identify the best methods to deliver high quality care (AHIMA 2017). To support this trend, the Health Information and Management Systems Society (HIMSS), a not-for-profit organization, focuses on improved health through information and technology. They started a training module serving clinical and non-clinical professionals for the purpose of covering major components of the analytics process (HIMSS 2018). While this training module is limited in content and methodology (not technical nor statistical), it strongly supports the need for interdisciplinary and skilled professionals to address the growing analytical needs in the domain. Developing the appropriate curricular material can potentially strengthen the skills needed by the next-generation of healthcare analytic experts. However, curricular material as well as pedagogical research are not widely available in business analytics programs.

To address this void, we developed an interdisciplinary healthcare analytics course to interconnect analytics skills with domain knowledge and decision making. Moreover, it is an elective course in the Business Analytics Masters Degree Program hosted by the Department of Computer Information Systems within the College of Business at Quinnipiac University.

Using a contextual active learning approach, we developed seven curricular modules for this course. Such an approach emphasizes active and contextual learning both in the course design and delivery. We present an experiential perspective and provide an account of our practical experience gained from teaching the 1.0
version of this course and discuss the implications for healthcare analytic training in the business school context.

In the following sections, we present a brief exposition of the healthcare analytics and active learning pedagogy. We then discuss our experience with teaching this course, and present the pedagogy, approach, course content, and software application tools. We also describe different course assignments and exercises to support a contextual active learning approach. We conclude with a discussion of lessons learned and survey scores for the associated course learning outcomes.

Related Work

In this section, we review existing healthcare analytics research on the use of active learning in teaching business courses. These two areas have direct relevance to our development of this healthcare analytics course.

Overview of Healthcare Analytics

As healthcare data proliferates, there is an urgent need for additional knowledge and skills to analyze this deluge of data and turn it into competitive advantage (Cortada et al. 2012). Healthcare analytics can provide organizations the ability to use their data to improve quality of care, increase financial efficiency and operational effectiveness (Raghupathi and Raghupathi 2014). Healthcare analytics refers to the use of tools (statistical, contextual, quantitative, predictive, etc.) for the purpose of providing actionable items for better decision making (Kankanhalli et al. 2016). The domain offers a panoramic view of the healthcare data and thus provides different stakeholders the ability to go beyond improving profits and reducing waste, to enable epidemic predictions, disease mitigation and cure, and quality of life improvements (Lin et al. 2017).

Raghupathi and Raghupathi (2013) describe a healthcare analytics four-stage model. The first stage, descriptive analytics, consists of categorizing and aggregating the data in order to understand past and current health care decisions. In healthcare, descriptive analytics is useful in answering questions such as: How many patients were treated? Which type of medical conditions were predominant? And what was the revenue generated by facility last quarter? Predictive analytics include “empirical methods (statistical and other) that generate data predictions as well as methods for assessing predictive power” (Shmueli and Koppius 2011, p. 553). In other words, it examines historical health data, detects patterns and then extrapolates these relationships to predict future outcomes. In predictive analytics, a health professional might seek to predict the type of patients who will respond to a given drug, patients who are most likely to have a medical condition (e.g., heart attack), anticipated costs, or predict medication failures. Unlike descriptive analytics, predictive analytics uses more advanced techniques and methods, such as data mining. The third stage of analytics is prescriptive. It uses medical and healthcare knowledge to supplement the outcomes of descriptive and predictive analytics to finalize a decision when more than one choice is available. Finally, discovery analytics utilizes “knowledge about knowledge, or wisdom, to discover new drugs (drug discovery), previously unknown diseases, alternative treatments” (Raghupathi and Raghupathi 2013, p. 4). It is worth noting that while the tools and methods are different in descriptive, predictive, prescriptive, and discovery analytics, many applications involve all four approaches.

Active Learning Approach

Traditional lecture format or lecturing has been the principal method of teaching dating back to 900 years ago when universities were founded in Western Europe (Brockliss 1996). The traditional format has been labelled as a passive methodology because students simply receive information from the instructor (Pinder 2013). Active learning, often contrasted to the traditional lecture format, has been defined as methods and activities where students engage in the learning process through problem solving and higher-order thinking (Prince 2004). In other words, active learning promotes “instructional activities involving students in doing things and thinking about what they are doing” (Bonwell and Eison 1991, p. 3). The active learning approach has received considerable attention and its effectiveness has been supported in diverse domains, including software development (Roussev and Rousseva 2004), operation management (Hill and Baker 2016), marketing (Laverie 2006), and engineering (Prince 2004).
Despite current studies advocating the importance of context when promoting active learning (Berkhout et al. 2018; Chung 2017), research on the development and delivery using contextual and active learning approaches is not widely available (Chung 2015). Our pedagogical approach seeks to address this need.

Course Pedagogic Approach and Learning Goals

A Contextual Active Learning Approach

In this study, we developed a contextual active learning approach based upon the active learning principles and contextual course components for the purpose of developing curricular modules for an online healthcare analytics course. A curricular module is an organized collection of objectives, activities, assignments and discussion spanning over a week. This is a graduate online intensive seven-week course. Our contextual active approach is based on the following key components: (1) Active learning – designing each module to include activities and lessons to promote active participations of learners. (2) Contextual development – modules are developed using healthcare as the focus and context for all activities for the purpose of relating skills and analysis to actual outcomes (3) Online adaptation – the course is delivered online in an unsynchronized mode and therefore the development adapts to online delivery intended to maximize learning.

Course Content and Learning Objectives

In this section, we present the course curricular modules in a format that aligns with the contextual active learning approach introduced above. This is an elective course offered online for graduate students majoring in Business Analytics. This course was offered online for the first time in Fall 2017 and includes the following learning objectives (LO):

1. Define data sources and key uses for healthcare analytics (LO1).
2. Compare analytical methods and tools used to analyze healthcare data (LO2).
3. Apply analytical methods and tools to solve selected healthcare problems (LO3).

While the course has a statistics pre-requisite, most students already had taken other quantitative courses, including predictive analytics and data mining, which are required courses in the business analytics program. Therefore, the context of healthcare is one of the attractive features of this course. The concepts are focused on (and emphasize) solving problems and/or creating value for healthcare professionals. These modules span seven major topics of the course (See Table 1). Each module objective is aligned with the course learning objectives noted above.

<table>
<thead>
<tr>
<th>Module and Topic</th>
<th>Contextualization &amp; Active Learning</th>
</tr>
</thead>
</table>
| **Module 1: Analytics in Healthcare** | **Module Discussion:** Discuss an example of a healthcare organization in using healthcare analytics to save lives, reduce cost, increase profit or other competitive advantages (LO1)  
**Data Analytics Tool:** Register for IBM Watson Analytics Account (LO3) |
| • Summarize the current state of data analytics in healthcare (LO1)  
• Classify types of Data Sources in Healthcare (LO1) | |
| **Module 2: Healthcare Data sources and Basic Analytics** | **Module Discussion:** Discuss the type of healthcare dataset(s) used in the study (LO1)  
**Descriptive Analytics:** Analyze healthcare security breaches using Watson analytics (LO3) |
| • Examine Electronic Healthcare Records (LO1)  
• Compare biomedical image data sources and analysis (LO1)  
• Examine sensor data used in medical informatics (LO1)  
• Discuss types of biomedical signals (LO3) | |
Module 3: Clinical Text and Social Media Data in Healthcare Analytics
- Analyze the role of Natural Language Processing (NLP) in healthcare (LO2)
- Examine social media and analysis for healthcare analytics (LO2)

Module 4: Healthcare Analytics Tools and Methods
- Examine clinical prediction models (LO2)
- Categorize temporal data mining for healthcare data (LO2)
- Discuss visual analytics for healthcare (LO2)

Module 5: Applications of Healthcare Analytics
- Examine data analytics for fraud detection in healthcare (LO1)
- Discuss data analytics for pharmaceutical discoveries (LO1)

Module 6: Challenges of Healthcare Analytics
- Discuss privacy and security issues related to healthcare analytics (LO1)
- Compare privacy preserving publishing methods (LO2)

Module 7: Healthcare Analytics – What’s Next
- Examine trends in healthcare analytics (LO1)

Module Discussion: How to locate relevant healthcare dataset and then analyze it (LO3)

Data Visualization: Analyze a dataset of interest (deeper look into descriptive analytics) (LO3)

Module Discussion: Engage in critical thinking activities related to term project (LO3)

Predictive Analytics: Use IBM Watson Analytics to find top drivers and other factors (LO3)

Module Discussion: Writing across the curriculum activity – how to communicate your descriptive and predictive and analytics findings? (LO3)

– Identify current healthcare analytics topic and match dataset for analysis (term project) (LO3)

Module Discussion: Discuss concerns/challenges (e.g., ethical, legal, data governance, silos, privacy, security...) as a result of the adoption of big data in healthcare (LO1)

Prescriptive Analytics: Analyze healthcare data using healthcare analytics techniques (beyond descriptive and predictive into prescriptive analytics) (LO3)

Module Discussion: Discuss career opportunities in healthcare analytics (LO1)

Assessment: understanding of key concepts related to healthcare analytics (LO1, LO2)

Table 1. Course Modules and Contextualization

Platform and Analytic Techniques Overview

The big data analytic platform employs a plethora of techniques and tools to handle the large amount of data. Recent studies have shown that students are discouraged when multiple platforms are used within a single analytics course. Although it increases their breadth of software use, it decreases the depth of acquired skills within each application (Asamoah et al. 2017). As such, we embraced one major tool for this course to allow in-depth mastery of the skills leveraged through the tool.

The choice of tools, skills and content was based on feedback from industry and academia experts. Since advanced statistical methods are taught and used in other courses in the business analytics program, a focus on the contextual - Healthcare Analytics - was emphasized rather than learning a major new and or complex analytics technique. We selected IBM Watson Analytics as our data analytics tool for the following reasons: It allows users to 1) execute both descriptive and predictive analytics, 2) create dashboards and infographics, 3) discover relationships and test correlations, and 4) is cloud based and therefore has no/limited hardware constraints (https://www.ibm.com/watson-analytics). Figure 1 provides an illustration of the various discovery capabilities as suggested by Watson. It also allows healthcare professionals to type their own inquiry (in the form of questions) once a dataset is uploaded. The lower section shows capabilities of testing relationships or running some predictive analysis.
Contextual Active Learning Assignments

We used a variety of assignments, discussion themes, and a term project that align with the course learning objectives. We present three sample course assignments that covered multiple concepts. For each assignment, we present the learning objectives, description and how it aligns with our contextual active learning approach.

Sample Course Assignment 1: Analyze Medical Data Breaches

- Learning Objectives:
  - Perform Extract-Transform-Load (ETL) of healthcare data dealing with breaches
  - Apply IBM Watson Analytics to analyze healthcare data
  - Perform descriptive analytics to answer the “What” questions

- Assignment Description:
The Privacy Rights Clearinghouse (PRC) Chronology of Data Breaches, accessible via http://www.privacyrights.org/data-breach, is a nonprofit corporation in California that was established in 1992. PRC keeps up-to-date information of data breaches across all industries and the government within the US. PRC aims to provide timely and historical information on data breaches. PRC reported more than ten billion records breached from over 7900 data breaches since 2005. Students are asked to apply analytical skills that will produce measurable insights from historical performance data that can be transformed into actionable insights. Students will engage in two major analytic activities: (1) extract, transform and load (ETL) the data; (2) create visualization graphs using business analytics tools (e.g., IBM Watson Analytics). In this assignment, students will focus on descriptive analysis and will be answering the “What” questions in order to provide a view of both current and historical results. Descriptive analytics tells the business how it is performing and help identify key issues in their current performances.

- Contextual Active Learning Focus
The students were required to extract publicly available dataset through PRC. They purposely selected medical/healthcare data for analysis since healthcare is the context of this course. With a focus on data breaches from healthcare, they found that these breaches account for almost half of the total reported breaches across all industries. Healthcare data breaches account for about 4000 cases of breaches out of 8300 cases across all industries, and impact more than 228 million patients (PRC 2018). The analysis activities using the contextual active learning approach is forcing students to engage in not only creating visualization charts but also thinking about the extent of the findings and engaging in what they are doing. The engagement and feedback were happening through the required weekly discussions.
Sample Course Assignment 2: Identify healthcare topic, dataset, perform analytics, and communicate findings

- **Learning Objectives:**
  - Identify a healthcare topic of interest and matching dataset for analysis
  - Analyze healthcare data using healthcare analytics techniques
  - Discuss and communicate findings in writing

- **Assignment Description:**
  Student will select a healthcare topic of his/her own choosing along with a dataset matching the topic selected. This requires researching and selecting an appropriate dataset (i.e., being able to explain or predict a phenomenon), preparing the data, analyzing the data, creating appropriate graphs, and telling about the analysis in a research-oriented format (introduction, literature review, methodology and dataset, findings, discussion and conclusion).

- **Contextual Active Learning Focus**
  Early in the semester, several students expressed their desire to use their own anonymized dataset available through their businesses. While the module included one source as a possible venue for datasets, many students researched and shared healthcare public dataset sites such as healthcare data.gov, public domain Medicare datasets, City of Chicago data portal (HHS data), and Statista. Such flexibility and engagement in the learning process allowed students to not only collaborate within their small group but engage beyond the requirements of the project. For example, some students added additional datasets (such as census data, weather data) to further their analysis. The grading was based on criteria that further supported more interactions among each group. Criteria included how interesting and challenging the student's project was (topic and dataset), the quality of analysis and writing, and peer evaluations.

Sample Assignment 3: Discuss Healthcare Analytics Issues

- **Learning Objective:**
  - Identify current healthcare analytics issues for analysis
  - Discuss implication and illustrate with current events/publications

- **Assignment Description:**
  The adoption of big data in healthcare increases security and patient privacy concerns. Write one short paragraph describing concerns/challenges (e.g., privacy, security, ethical, legal, data governance, silos) as a result of big data adoption in healthcare.

- **Contextual Active Learning Focus**
  After watching an entertaining video of ordering pizza (Dedots 2006), students quickly identified some implications and challenges of data analytics. The video portrays the dismay of a customer who was placing a pizza order. Not only information about his phone, national ID, work, and recent travel booking was available to the clerk, but also his health data (high cholesterol and high blood pressure) that caused a premium cost to his pizza, Tying a simple pizza order to health complications puts in perspective the power and possible concerns of analytics. To encourage engagement and discussions about the issues and the context of analytics in healthcare, students were required critique their peers.
Designing a Healthcare Data Analytics Course

Discussion

In this section, we discuss students’ assessment of this course. We also discuss the lessons learned and offer recommendations to enhance future deployments.

Learning Assessment

The course evaluations have shown that students gave positive feedbacks of overall satisfaction with the course. We assessed students’ perceptions of the course impact and instructor effectiveness through a survey using a 5-point Likert scale. The results of the survey are provided in comparison to other courses within the College of Business. The students (n=10) rated the instructor the overall teaching ability at 4.6 out of 5.00 in comparison to the average of the school of business instructors rating at 4.0 out 5.00. They acknowledged that the instructor brings current ideas and emphasizes the intersection of healthcare and analytics and summarizes which is crucial in contextual settings as stated by one student in the survey:

“Very helpful, intriguing, enlightening, and extremely relevant to the real world.”

In alignment with an active learning approach, the course included a variety of individual and group activities that were not only engaging but yielding maximum learning in technical, analytical and team work skills. For instance, one student stated:

“I felt group activities were engaging and required group members to utilize relevant technical and social skills to collaborate.”

Another student commented:

“I think this class has a healthy balance of individual work assignments and team assignments. During my early days in the Business Analytics program, many of the courses were too heavily centered on group work and there was no individual accountability. The classes are evolving and creating an environment that facilitates maximum learning.”

Though not captured directly by the survey, the open-ended questions provided more feedback on the selected analytics platform. The course aims to create a balance between content and technology on one hand and between students from different background. Students who were already working in industry and from technical background felt a certain level of frustration when they could not manipulate certain functionalities within the selected software as stated by one student on the open questions:

“The material is well structured and thought out. The learning objectives are clear as are the summary materials she created. Dr. Parks puts a lot of effort into connecting with students which is much appreciated. Though it was a new topic area and one that is quite interesting, I do not think I enhanced my analytical skills through this class, particularly through the use of Watson Analytics. I am generally disappointed by how limited and underdeveloped Watson is and do not consider it an analytical tool that can be added to my toolbox.”

In addition to the course evaluation completed by students in this first round, we are currently considering additional assessment to evaluate the course and modules. New evaluations will include different methods, such as qualitative data gathering from industry stakeholders, and assessment of instructional design and delivery. Incorporating feedback from students, employers and educational designers may further ensure the interdisciplinary goals of this course while ensuring effectiveness of learning.

Lessons Learned

We present some of the key challenges we encountered throughout this course in the form of lessons learned:

Textbook or No Textbook

While designing and developing this course, we were able to put together the topics to cover based on the authors’ industry experiences, research in healthcare analytics and assistance from other experts in industry
and education. Once the topics were identified, we selected a textbook. The first challenge we debated for months whether to assign or not assign a textbook for this interdisciplinary course. We erred on the side of assigning a book simply because the course is delivered online, and it provides a sense of structure to the students. Selecting the appropriate book for any course is not trivial. However, when the course is interdisciplinary and assumes different background for students, the task becomes very daunting. We attended vendors’ booths at Information Systems and Health Informatics conferences – Americas' Conference on Information Systems (AMCIS), the Information and Management Systems Society (HIMSS), and the Conference on Information Systems and Computing Education (EDSIG) – and approached professors and editors about health analytics books. This task was a very challenging because 1) there is a limited number of “realtime” books that covers data analytics in healthcare and 2) among the few we found, each book was taking a completely different perspective. The first book looked promising based on the description. After we ordered and reviewed the book, it was almost anecdotal. The second one was business/managerial in nature, and we almost went that route before realizing it barely discussed healthcare dataset sets or different analytical methodologies. Finally, we chose a book edited by Chandan Reddy who is a Computer Science professor and Charu Aggarwal who works at IBM Research Center (Reddy and Aggarwal 2015). Our initial reaction was not to use the book because it may seem geared toward computer science degree students especially with some sections getting into the algorithms behind the different methods. Then, we closely assessed the content of each chapter, and noticed the variety of academic and industry background of the authors writing each chapter and thus providing the interdisciplinary features that met the course learning objectives. For example, Chapter 1 is written by a computer scientist and IBM researcher while chapter 3 is written by employees of the biomedical image analytics lab at GE Global Researcher. Chapter 4 is written by employees at the IBM Watson research lab and College of Medicine researchers while chapter 12 is written by sciences researchers from the School of Information Library Science.

While the book touches on the major learning objective set for this course, we still have some reservations. Some chapters are written by computer scientists where much attention is focused on the algorithms which is outside the scope of the course and our graduate business students. This was alleviated by making clear to students that while the details are great, they are not part of this course and students are not expected to learn nor memorize this material. The other reservation pertains to the book lacking coverage of the social and behavioral aspects of data analytics, such as privacy, patient engagement and ethics topics, but those were easily covered in the course through relevant research articles.

Software Platform

The software used has a significant impact on the overall class. As tempting as it may be to expose the students to different analytics platforms, we learned from other data analytics classes we taught not to introduce many data analytics platforms during one course offering. This is also recommended by other researchers who state that using many applications enriches the students’ experiences but also has major drawbacks especially with the width of coverage and the lack of depth needed to perform the analysis (Asamoah et al. 2017). As this is an elective course (with statistics as prerequisites), we did not have to worry about covering the statistical or analytical foundations of business analytics. Therefore, our focus was on a 1) widely used software in industry especially in healthcare, 2) moderately challenging but not overwhelming software so it does not take away from the contextual learning of healthcare, and 3) finally offering both descriptive and predictive capabilities. We selected IBM Watson Analytics as our platform to meet the above criteria.
Interdisciplinary Backgrounds and Flexibility

We assessed the academic backgrounds and professional experiences of students taking this course. A total of 21 were enrolled in this elective business analytics course with academic background varying from management information systems to fine arts and actuaries. All students who completed this information (19 students out of 21) are currently working in a variety of industries (healthcare, insurance, government, manufacturing ...). Despite the disparities in their academic backgrounds and work experiences, all of the students reported on the use and of the need of analytics in their current jobs. Such in disciplinary backgrounds provided a very rich platform for discussions for each module and allowed students to apply their learning to their work environment and engage beyond the requirements of the course. This also dictated a flexible structure to move from an assigned term project to a project allowing students to use their own dataset.

Contextual Active Learning

The combination of a data analytics and healthcare attracted students to this graduate elective course. Based on utilizing an active learning approach, we introduced several learning objectives for each module and align each activity (e.g., assignments, readings, discussions …) with the learning objectives (see Table 1). With the course being contextualized in healthcare and half of the students working in healthcare or related field (healthcare insurance), the assignments especially the term project and the weekly discussions took a much deeper undertaking both at the data analysis and discussions activities.

Online vs. On Ground

This course was offered online which is different from on ground offerings. The development of the course did not make any assumptions about knowledge or skills. For instance, we set a webinar with the librarian to offer a session on how to search for discussion articles, use different databases and search engines. In an on ground class, the entire activities would have a show of hands to check if they know how to search for scholarly articles. Sensitive to different backgrounds, skills are necessary to create a conducive learning environment.

Healthcare Domain

Despite the first module being dedicated to analytics in the context of healthcare, students worked on interpreting data/models without a clear competency on the sector of healthcare itself. The absence of a clear and dedicated overview of general healthcare domain knowledge could be easily alleviated by providing such background to all students in the first module.

Conclusion

This research highlights the need to integrate context-specific learning of analytics through a specialized analytics course. Rising concerns in healthcare data analysis call for skilled professionals to appropriately analyze healthcare data to improve care, predict epidemics, and reduce preventable deaths. However, curricular material that intersects with healthcare and data analytics are not widely available. This research utilized a contextual approach to creating curricular modules for use in healthcare analytics. The results should contribute to addressing needs for active-learning and contextualized learning in healthcare data analytics, developing skilled data analytics for healthcare, and providing empirical results on a healthcare analytics course implementation.

REFERENCES


