2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #32932

WIP: Student Training in Data Analytics Approaches for Bioprocessing Through Co-Curricular Activities

Dr. Maryam Mobed-Miremadi, Santa Clara University

Dr. Maryam Mobed-Miremadi is a Senior Lecturer at Santa Clara University. She is a chemical engineer (B Eng, M Eng, McGill University Canada) by training with an emphasis on transport and microencapsulation technologies for bioengineering applications. She graduated from McGill University with a Ph.D. in Chemical Engineering having conducted her research at the Artificial Cells and Organs Research Centre. Following graduate school she spent 12 years in the biomedical industry, more specifically in Diagnostics (Abaxis), Genomics (Agilent Technologies, Wafergen) and Biomedical Devices (Boston Scientific). After serving as Kordestani's Chair in Bioengineering at San Jose State University, she joined the department of Bioengineering at Santa Clara University.

Navid Shaghaghi, Santa Clara University

Navid Shaghaghi is a lecturer and researcher in the department of Mathematics and Computer Science in the College of Arts and Sciences, the departments of Bioengineering and Computer Science & Engineering in the School of Engineering, and the department of Information Systems and Analytics in the School of Business at Santa Clara University. He directs the Ethical, Pragmatic, and Intelligent Computing and Creative, Augmented, and Virtual Environments research laboratories, and is a faculty fellow at the Frugal Innovation Hub and the BioInnovation and Design Lab at the university.

Gangshu Cai, Santa Clara University

Dr. Cai is a full professor and Department Co-Chair of Department of Information Systems and Analytics, Leavey School of Business. He is the former Faculty Director of Graduate Business Program and founder of undergraduate Minor in Business Analytics. Dr. Cai is an Associate Editor of Decision Sciences Journal and a Senior Editor of Production and Operations Management Journal.

Dr. Prashanth Asuri, Santa Clara University

Dr. Prashanth Asuri joined the Bioengineering faculty at Santa Clara University in Fall 2011 after holding research positions in academia and industry. Dr. Asuri earned his Ph.D. in Chemical and Biological Engineering from Rensselaer Polytechnic Institute and M.B.A. with a focus on Leading Innovative Organizations from Santa Clara University. His research efforts focus on the design and development of simple in vitro platforms to understand complex in vivo phenomena. He teaches undergraduate and graduate courses on topics related to tissue engineering, biomaterials, and bioinnovation. Dr. Asuri also serves as the Director of SCU's BioInnovation and Design Lab that partners with industry to empower students to discover, innovate, and address complex challenges within healthcare.

Work in Progress: Student Training in Data Analytics Approaches for Bioprocessing through Co-curricular Activities

Introduction

The rapidly growing biologics category now makes up more than 30 percent of the overall biopharmaceutical market, with the majority of manufacturing production focused on protein-based drugs [1]. As a new wave of clinical trials progress, a wider variety of new gene, cell, exosome, and recombinant protein therapies are moving closer to commercialization and manufacture [2]. The complexity and cost of these biologics, as well as the unique manufacturing requirements and patient markets for each therapy, present significant engineering, business operations, and manufacturability challenges with many additional variables for consideration to automate biopharmaceutical manufacturing. Advanced data science techniques, which have made a significant impact on biopharmaceutical discovery research in recent years, also offer the potential to transform bioprocessing platform automation [3]. These techniques are well suited to intricate technical challenges, like bioprocessing for cell culture, gene therapy, and cell therapy, that present a large number of variables, the potential to generate large trainable datasets, and complex processes. To deliver on the promise of this coming digital transformation, however, significant investments are needed in workforce training [4]. Most activities that prepare manufacturing technicians via training on emerging bioprocesses are at community colleges. Additionally, equipment manufacturers or university consortia offer training solutions (in-person and online) to introduce almost all emerging bioprocessing techniques and equipment. However, fewer initiatives focused on training incumbent engineers on data-driven bioprocessing techniques, especially at the bioprocessing platform and manufacturing facility levels. We hope to address this gap in extant education and workforce training solutions to support the digital transformation in bioprocessing. Over the long-term, we envision building toward an industry-academic consortium that supports multiple training and education needs related to data analytics across biopharmaceutical manufacturing hubs. As a first-step, our project identified a subset of achievable near-term tasks and objectives needed to develop and pilot an introductory series of active-learning boot camps designed to upskill incumbent employees serving in established biomanufacturing scientific and engineering roles (e.g. not trained data scientists) to utilize multivariate analysis tools and extract value from complex data sets. Our proposed solution is differentiated by its active collaboration with industry, and commitment to mentored employee knowledge integration into corporate workflows.

Objectives and Value Propositions

We propose the development of a workforce development collaboration that (1) trains incumbent workforce to make data-informed decisions in manufacturing platform automation, (2) develops applied simulations and decision-making models that supplement workshop training, (3) prepares and disseminates relevant industry cases, and (4) generates new datasets that can be used by researchers to build, train, and test predictive analytical techniques that advance the state of the art. Through these efforts, we aim to address an emerging workforce readiness gap in data mining and analysis for bioprocess development that hinders the digital transformation of the biopharma manufacturing industry and resultant gains in quality, productivity, cost-efficiency, and speed-to-

market. In 2020, we successfully submitted a proposal to the National Institute for Innovation in Manufacturing Biopharmaceuticals (NIIMBL) to support the development of co-curricular programs (to be delivered in the form of bootcamps) that provide training in data engineering and analytics approaches for bioprocessing and manufacturing. While the proposed primary outputs of the project are to develop training programs for upskilling and cross-training incumbent engineers, we postulate that the content may also be leveraged for collegiate pipeline development. Such bootcamps would complement existing university programs that already have embedded multivariate analysis content into their course offerings, albeit in stand-alone formats that may not be easily integrated into cross-disciplinary training efforts.

To determine the relevance of the subject matter at various educational and professional levels and to identify specific knowledge and skills gaps that may exist for different target audiences, evaluations will be differentiated between incumbent worker success and collegiate pipeline success. Success means that student participants report that the experience has informed their career path and increased their awareness of opportunities of data analytics in the biopharmaceutical industry. Additionally, they report interest in enrolling in relevant courses to gain depth and proficiency to the point where their work products can contribute to faculty publications in the field. In the future, as we expand our efforts to offer certificate programs, or other credentials, we will develop assessment tools to confirm effectiveness of the learning content and delivery.

Current Efforts and Next Steps

We first consulted with biopharmaceutical companies with different bioprocessing requirements and life science associations, and identified a set of learning objectives to support upskilling of incumbent employees serving in established engineering roles in bioprocessing to utilize multivariate analysis tools and extract value from complex data sets. The primary objectives of the bootcamps are to provide an understanding of how bioprocessing data may be utilized to develop reliable decision-making models, and how data analytics may be applied to improve bioprocessing and biomanufacturing. We expect that at the end of the bootcamps, divided into two sessions – Introductory and Advanced, the attendees will be able to: (1) apply statistical thinking and methods to identify and diagnose problems and make informed decisions; (2) generate robust decision-making models using biopharmaceutical manufacturing data from upstream and/or downstream manufacturing operations; and (3) apply software-based tools and scripts (e.g. python) to solve the model(s) and interpret the results. Learning outcomes were common to introductory and advanced bootcamps and designed to engage attendees with different levels of proficiencies.[a] Please note that assessments of outcomes were not performed for the pilot and will be included in future efforts.

We delivered the first bootcamp (focused on introductory concepts) in Fall 2020 simultaneously to cohorts of incumbent employees from our NIIMBL industry partner (23 attendees) and non-professional collegiate students at our university (27 attendees) from different relevant departments (e.g., bioengineering, biology, computer science, etc.), with each attendee receiving close to ten hours of instruction and virtual support. The bootcamp sessions covered the application of various statistical analyses (e.g. regression, hypothesis testing, ANOVA) to upstream and downstream unit

[a] Additional details on the bootcamp learning outcomes, topics and the survey results have not been included here due to the space limitations and will be presented at the conference presentation.

operations (bioreactor, flocculation, and chromatography).[a] These topics were chosen to illustrate problem solving rigor applied to bioprocess operations frequently encountered in the industry. By testing content dissemination to university students (alongside incumbent employees), we hoped to validate usefulness of the content in collegiate pipeline development. Participant evaluations were leveraged to guide continuous quality improvements, with special attention given to skills and knowledge areas with which the different participants are struggling. The surveys indicated that participants enjoyed the "interdisciplinary and applied nature" of the bootcamps, "with a good blend of bioscience, statistics and python modeling" and content "explicitly tied to practical applications". Also, a significant subset of the attendees (>75%) confirmed that they would recommend the bootcamp to their colleagues.[a] The survey results also indicated that both cohorts did not have any difficulty following the concepts presented in the bootcamp, thus suggesting that there is not a compelling need to adopt separate content development and dissemination strategies for incumbent workforce and college students. Furthermore, main topics in the bootcamp (i.e. statistics, programming and bioprocessing) are taught as separate courses, and college students attending the workshop may have taken any combination of these classes, and the content was adjusted to basic proficiency level. However, it is important to note that a few attendees expressed a need to be better prepared for the python modeling sections, something we plan to address by creating introductory primers on the use of python for data analytics in bioprocessing. These findings have to be confirmed for bootcamps that provide an advanced understanding of how analytics can be applied for bioprocessing; however, it is encouraging to see that bootcamps developed for incumbent worker training may be leveraged for collegiate pipeline development.

Our next steps are to develop cases and companion data sets that will be disseminated to NIIMBL members and support development of advanced training modules. We will also deliberate upon ways to adapt content from the bootcamps into existing or new curriculum through course activities including assignments, group discussions, and student presentations on applications of data analytics in bioprocessing. We will sustain and scale our impact through the development of a consortium of corporations, industry associations, and universities that collaborate to develop and disseminate training content, with a secondary goal of adapting content for credentials (e.g. badge, certificate) offered remotely or in hybrid form. We plan to share the program at future conferences as a model for sustained engagement with industry that could be replicated at other schools. We also plan to pilot a collaborative collegiate summer institute for data analytics in biomanufacturing that offers applied learning opportunities for rising seniors and graduate students from traditionally underrepresented socioeconomic backgrounds (e.g. first-generation college students).

References

1. M. Kesik-Brodacka, "Progress in biopharmaceutical development," *Biotechnol. Appl. Biochem.*, 65(3), pp. 306–322, May 2018.
2. S. Zobel-Roos, et al. "Accelerating biologics manufacturing by modeling or: is approval under the QbD and PAT approaches demanded by authorities acceptable without a digital-twin," *Processes*, 7(2), pp. 94, Jan. 2019.
3. Fleming, N., "How artificial intelligence is changing drug discovery," *Nature*, 557 (7706), pp. S55–S57, May 2018.
4. R. F. Wright, "The future of biopharma hinges on the ability to manufacture," *Life Science Leader*, 20-50. Dec. 2018. [Online]. Available at: https://www.lifescienceleader.com/