2023

Development of a Data Science Curriculum for an Engineering Technology Program

Salih Sarp
Old Dominion University, ssarp@odu.edu

Murat Kuzlu
Old Dominion University, mkuzlu@odu.edu

Otilia Popescu
Old Dominion University, opopescu@odu.edu

Vukica M. Jovanovic
Old Dominion University, v2jovano@odu.edu

Zafer Acar

Original Publication Citation

Sarp, S., Kuzlu, M., Popescu, O., Jovanovic, V. M., & Acar, Z. (2023, June). *Development of a data science curriculum for an engineering technology program* [Paper presentation]. 2023 ASEE Annual Conference & Exposition, Baltimore, Maryland. https://peer.asee.org/43109

Development of a Data Science Curriculum for an Engineering Technology Program

Salih Sarp, Old Dominion University

Salih Sarp is a Ph.D. student in the Electrical and Computer Engineering department at Old Dominion University, USA. Currently, he is developing AI applications and sensor fusion models. Previously, he received his BS degree in Electronics and Communication.

Dr. Murat Kuzlu, Old Dominion University

Murat Kuzlu joined the Engineering Technology Department at Old Dominion University (ODU) as an Assistant Professor in 2018. He received his B.Sc., M.Sc., and Ph.D. degrees in Electronics and Telecommunications Engineering. He worked as a senior researcher at TUBITAK (Scientific and Technological Research Council of Turkey) between 2006 and 2011. Before joining ODU, he was a Research Assistant Professor at Virginia Tech's Advanced Research Institute. His research interests include cyber-physical systems, smart cities, smart grids, artificial intelligence, and next-generation wireless networks.

Dr. Otilia Popescu, Old Dominion University

Dr. Otilia Popescu received the Engineering Diploma and M.S. degree from the Polytechnic Institute of Bucharest, Romania, and the PhD degree from Rutgers University, all in Electrical and Computer Engineering. Her research interests are in the general areas of communication systems, control theory, signal processing and engineering education. She is currently an Associate Professor in the Department of Engineering Technology, at Old Dominion University in Norfolk, Virginia, and serves as the Program Director for the Electrical Engineering Technology Program. In the past she has worked for the University of Texas at Dallas, University of Texas at San Antonio, Rutgers University, and Politehnica University of Bucharest. She is a senior member of the IEEE.

Dr. Vukica M. Jovanovic, Old Dominion University

Dr. Vukica Jovanovic is a Chair of Department of Engineering Technology and Associate Professor and Batten Endowed Fellow in Mechanical Engineering Technology Program. She holds a Ph.D. from Purdue University in Mechanical Engineering Technology, focuses on Digital Manufactur

Zafer Acar

Development of a Data Science Curriculum for an Engineering Technology Program

Abstract

Data science has gained the attention of various industries, educators, parents, and students thinking about their future careers. Statistics departments have traditionally offered data science courses for a long time. The main objective of these courses is to examine the fundamental concepts and theories. However, teaching data science courses has also expanded to other disciplines due to the vast amount of data being collected by numerous modern applications. Also, someone needs to learn how to collect and process data, especially from industrial devices, because of the recent development of Internet of Things (IoT) technologies. Hence, integrating data science into the curricula of different engineering branches becomes a matter of relating the statistics background to the specific discipline. There are several reasons for this transition. Firstly, as the increased computational power and massive availability of the data make the use of statistical theories possible in more engineering applications, there is a growing need for engineering students to build knowledge in data science concepts. Secondly, the wide availability of libraries and models allows for the implementation of diverse solutions to engineering problems. This paper will discuss introducing a new data science curriculum in an Engineering Technology (ET) program with a focus on Electrical Engineering Technology (EET) program.

Introduction

A typical data science curriculum covers various topics such as data processing, feature engineering, regression, classification, and natural language processing. However, it does not cover concepts related to data acquisition and processing from the standpoint of technology and hardware. In the last decades, data-driven models have significantly affected almost every industry. There are various courses across different programs and curricula in the nation that focus on introducing data science topics. However, a complete Engineering Technology (ET) curriculum that would embed data science topics related to data retrieval, instrumentation, and applications related to hardware equipment has not been developed yet, and emerging technologies have not watched the pace in the typical engineering technology curriculum that is focusing on electronics and data acquisition.

Data science has dramatically expanded with high demand in many industries, from energy, healthcare, finance, manufacturing, and many more[1]. As a result, there is a growing call for data scientists and engineers who can work with large amounts of data and extract meaningful insights to support decision-making in all industries, and in particular in manufacturing and processing engineering applications. To meet this demand, the development of a data science curriculum for an engineering technology program is critical to provide the current and future workforce able to respond to new challenges. Also, it is making engineering technology students better prepared for emerging technology trends since data acquisition and processing are related to modern data science tools and how they are used to make decisions about controlling the hardware. This curriculum should be designed to equip students with the necessary skills and knowledge to work with data, including statistical analysis, introduction to artificial intelligence tools, introduction to machine learning concepts, introduction to data visualization tools, and more[2].

Data science will be useful for the entire spectrum of engineering technology students. It can be a separate bachelor-level engineering technology department or graduate-level program, such as MS and Ph.D. in Engineering Technology, as well. Furthermore, data science has incorporated various techniques from the field of signal processing, encompassing signal analysis, neural networks, deep learning, IoT, and others. Some experts view machine learning as an extension of statistical signal-processing methods. That's why our focus in this paper is on developing and integrating of data science curriculum for Electrical Engineering Technology (EET).

A problem that we have identified is that the traditional Electrical Engineering Technology curriculum does not usually cover these topics and that there is a need for integrating more statistical methods in engineering technology courses[3]. However, the industry is changing at a much faster pace, and understanding of basics of these concepts will assist our graduates in working on their applied engineering jobs in the industry. One such application is the use of artificial intelligence in the training of vision sensors in the manufacturing industry to detect failures in industrial quality control and inspection. In the past, these tools were very difficult to implement but nowadays different manufacturers of industrial sensors are providing AI training tools to our graduates who are working as electrical engineers on the manufacturing floor designing manufacturing cell automation. AI tools have been integrated into machine vision applications[4] that are now, in a plug-and-play kind of manner, being integrated into daily engineering applications. Machine vision and machine learning technologies are now being integrated with automated material identification technologies that work with robots in IoT application setting[5]. Hence, the Electrical Engineering Technology (EET) curriculum should adapt to these new technology developments and also provide students with hands-on experience in applying data science techniques and tools with modern automation hardware and help them to develop the necessary skills to work with data.

Need for Integration of Data Science Skills within the Electrical Engineering Technology Curriculum

The main job definition of a data scientist occupation, according to the U.S. Department of Labor, is to be trained on how to use analytical tools and techniques to extract meaningful insights from data[6]. Recently, that job has gained significant attention due to the high salaries related to the occupation, increase in job opportunities, and comfortable working conditions[7]. However, someone needs to collect this data. In recent years, there has been an increasing interest in developing data science curricula and workforce development for many associate and bachelor's programs[7]. Hence, this is especially important for the field of engineering technology, in which many students continue their undergraduate education after completing their associate degrees.

For example, the National Science Foundation has launched several initiatives to engage more students outside of traditional data science programs. One of these initiatives is the ”Developing the 21st-century data science workforce” program[1]. There have also been several research studies focused on the development of data science curricula for engineering technology programs. For example, the authors in[8] explored integrating Data Science into a General Education Information Technology Course, implemented a course built around these ideas, and found that the course helps develop data-savvy in students. There are also several educational programs that offer data science courses for engineering technology students. A group of instructors in[9] provided a comprehensive guideline for undergraduate programs in data science. The team was comprised of 25 university instructors who primarily came from mathematics, statistics, and computer science disciplines and were mostly from various colleges and universities in the United States. These guidelines have been created with the intention of offering a framework for educational institutions that are either starting or updating their data science major programs.

Data science is usually integrated into engineering fields and a graduate level, and it covers a range of topics, including programming, statistics, data analysis, machine learning, and other related areas. These programs may also include hands-on projects, internships, and other experiential learning opportunities that can help students apply their knowledge and gain practical experience in the field[10]. However, if an industrial setting is already integrating data science skills in the day-to-day job of electrical engineering technology graduate, the curriculum has to adapt accordingly. By embedding an introduction to data science within the engineering technology program, undergraduate students will be able to develop a range of valuable skills in high demand in many industries. These skills include the ability to collect, organize, and analyze large amounts of data, as well as the ability to develop and apply machine learning algorithms and other advanced analytical techniques. These skills are highly sought-after by employers in many industrial settings in which many electrical engineering technology graduates work after graduation. Hence, the development of a data science curriculum for an engineering technology program may become imperative for preparing students for the future workforce. Along with the right skills and knowledge, engineering technology students will be well-positioned to tackle complex data-related problems in current and future engineering projects.

The Rational of Data Science Curriculum as part of EET

The Electrical Engineering Technology (EET) program at Old Dominion University started with offerings of traditional Computer Engineering Technology and Electrical Systems Technology. Recently, in response to the trends in the industry and the variations in the job markets, the program decided to divide the Electrical Systems concentration into four more specific areas of specialization closely related to the upper-division courses that were offered in the program. Currently, the program offers the following concentrations:

- Computer Engineering Technology - which combines hardware-focused core coursework with a package of Computer Science courses, which also fulfill the requirements for a

[1]https://beta.nsf.gov/science-matters/developing-21st-century-data-science-workforce

Computer Science minor.

- Communication Systems Technology - which involves specific courses that cover communication system principles and wireless communications, with the option to take additional courses in computer networks[11].
- Embedded Systems Technology - which focuses on courses in microprocessors/microcontrollers and microprocessor-based design.
- Mechatronics Engineering Technology - with a curriculum that includes courses on PLCs, microcontrollers, and power and communication systems, as well as a package of mechatronics-related courses in Mechanical Engineering Technology[12], which correspond to a minor in Mechanical Engineering Technology which includes topics related to the data acquisition and instrumentation[13].
- Power Systems Technology involves courses in power systems and machinery, alternative energies, and smart grids.

While all of these concentrations are built on core courses in electrical circuits, analog and digital electronics, and microprocessor applications, each of the tracks includes at least two to three courses specific to that concentration, these courses being at the 300-400 level, and would otherwise qualify as elective courses. The addition of a new Data Science concentration in the program would follow the same model to build on the core curriculum and add up to three data science-related courses. One such course could be Introduction to Data Science. An introductory data science course focuses on teaching students the fundamentals of the field. The course should cover specific topics, including data exploration and visualization, statistics and probability, machine learning, data cleaning and preparation, and visualization. For students that would take this course as an elective, the goal of the course will be to provide them with a solid foundation in the basics of data science. For students pursuing the Data Science concentration, this course will prepare them for more advanced study in the following course. The course focuses on giving students the skills and knowledge necessary to analyze and extract insights from data and to make informed decisions based on the insights they uncover. The current EET curriculum includes an Advanced Technical Analysis course that covers basic statics concepts and programming, which is a core course, and this course will provide the foundation for the Data Science concentration. Additionally, the current Instrumentation course, also a core course in the EET curriculum, covers principles of digital instrumentation devices used in process control systems, including sensors, actuators, signal conditioning circuits, and data acquisition boards, and students are learning to apply statistical methods to data acquisition and processing using Matlab and LabView software. This course will also be aligned with a Data Science curriculum.

In addition to the Bachelor of Science in Electrical Engineering Technology (EET), the program also offers a minor in EET. In the current format, this minor includes two 300-level required courses, and it gives the students the flexibility to choose two additional courses out of a variety of 300 and 400-level courses. These courses are within the EET elective courses group and are the courses that otherwise are specific to the concentrations offered in the program. The data science courses will add to the list of options for fulfilling the minor requirements in EET. This way, the minor is expected to attract a greater number of students from different fields, and the data science courses are expected to generate a lot of interest.

Data Science Course Descriptions and Contents

The data science courses, i.e., “Introduction to Data Science” and “Advanced Topics in Data Science,” will be offered as 400-level courses in the Electrical Engineering Technology (EET) program. These courses introduce the students to the fundamental principles and techniques in data science and machine learning for data collection, processing, analysis, visualization, and data-driven decisions. The introductory course will provide an overview of data science and machine learning and a comprehensive knowledge of their tools, along with hands-on activities. This course will also cover several course activities focused on the data science topics, such as introduction to data science and statistics with Python, data structures, data visualization, regression, classification, and clustering methods, dimensionality reduction, network analysis, computer vision with OpenCV, natural language processing.

The introductory course is designed for senior undergraduate students and only requires a prerequisite course covering the fundamentals of statistics and programming language, the Advanced Technical Analysis that was mentioned previously. Some background in calculus and linear algebra would also be helpful, but the concepts needed are reviewed in the course. For this course, no book is required, but a list of references is provided, and handouts based on these references are provided along with the selected journal and conference papers.

Having successfully completed the Introduction to Data Science course, the students will be able to:

- Understand the fundamental principles and techniques in data science and machine learning
- Collect data through different resources
- Clean and reshape collected datasets
- Understand exploratory data analysis tools for clustering and visualization
- Perform regression, classification, and clustering methods
- Apply dimensionality reduction tools
- Perform basic analysis of network data
- Perform natural language processing and sentiment analysis
- Evaluate the performance of machine learning methods

For course assessment and grading, the following instruments are used: two in-class tests and a final exam, random quizzes, and a term project. For the term project, students are required to prepare two drafts during the semester and a final report. The term project concludes with a final presentation.

Traditional data science courses generally focus on statistical analysis and the use of statistical software by emphasizing statistical modeling, hypothesis testing, and regression analysis. The curriculum often uses these tools to analyze and draw insights from structured data sets. In contrast, modern data science courses emphasize the practical skills required for working with big data, such as data cleaning, preprocessing, and feature engineering. Modern data science courses also often incorporate machine learning techniques, including deep learning and neural networks. In addition, modern data science courses use open-source tools and programming languages, such as Python and R, which are more widely used in industry and academia. Furthermore, modern data science courses often cover more advanced topics in data science, such as distributed computing, cloud computing, and data analytics.

The differences between traditional and modern data science courses show the changing nature of the field and the growing importance of practical skills for working with big data and using advanced techniques for data analysis.

Fig. 1 shows the fundamental and advanced topics in the data science curriculum. The fundamental topics cover several introductory course sub-topics, such as introduction to data science and statistics with Python, data visualization and information analysis, regression and classification methods, and deep learning. The advanced topics cover more complex contents for an advanced data science course, such as natural language processing, computer vision, recommendation systems, networking and graphical models, and AutoML, Big Data.

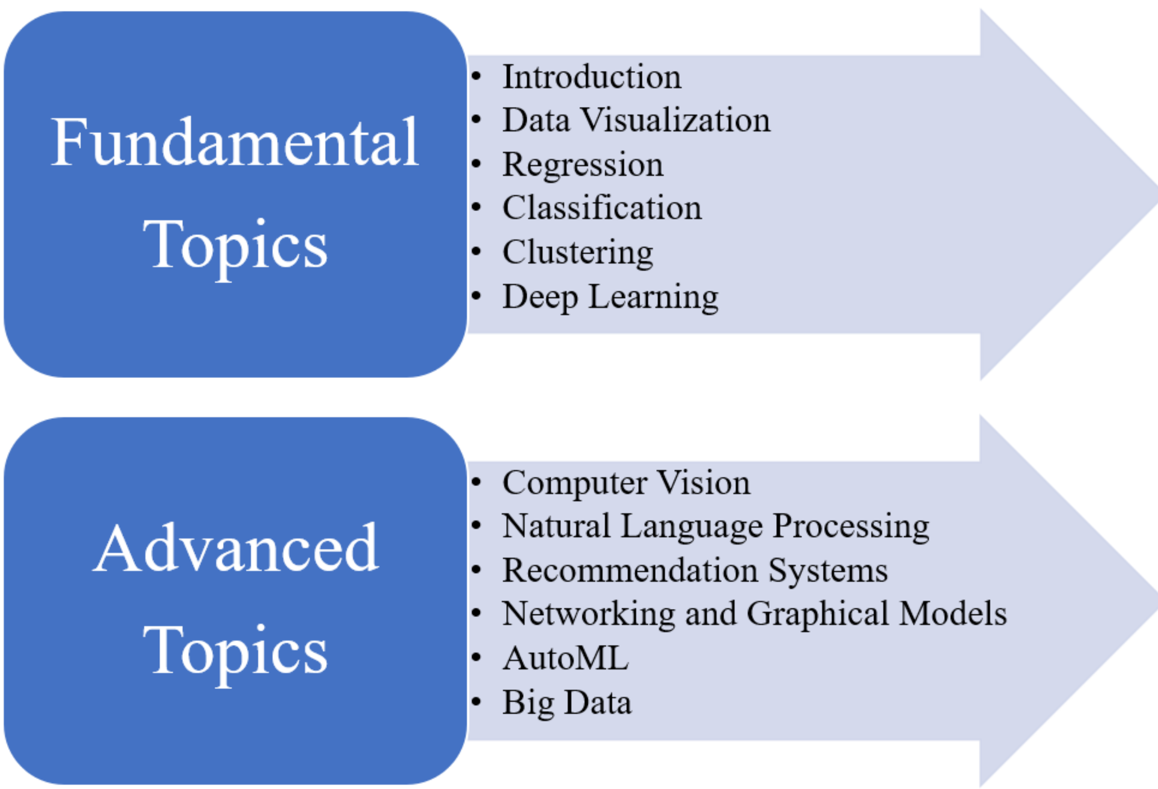

Figure 1: Fundamental and advanced topics in data science course

Fundamental topics in an introductory data science course provide a strong foundation in the basics of data science, each with a brief description is given as follows:

- Introduction to data science and statistics with Python: Learn the basics of the python programming environment, including fundamental python programming techniques and basic statistical analyses with Python.

- Data visualizations and information analysis: Perform data visualization in Python, and use various Python data visualization modules such as Matplotlib, Seaborn, Plotly, etc., and other such data visualization packages with different features for creating informative, customized, and appealing plots to present data most simply and effectively.
- Regression methods: Review and explain the difference between supervised learning and unsupervised learning methods, then build the most popular regression models, such as linear, polynomial, ridge, lasso, and gradient.
- Classification methods: Use pre-categorized training datasets, and predict a categorical outcome by performing classification methods in machine learning, such as Naïve Bayes, Decision Trees, Random Forest, and other techniques, as well as performance metrics, such as R-squared, MSE, RMSE, precision, recall, sensitivity, specificity, accuracy, AUC, ROC, along with gains and lift charts.
- Clustering methods: Investigate the clustering or cluster analysis methods in machine learning techniques for grouping the unlabelled dataset, such as K-Means, Mean-Shift, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN).
- Deep Learning: Introduce deep learning through Tensorflow, training neural network, and visualizing what a neural network has learned using TensorFlow Playground.

Each one of the advanced topics in the advanced data science course is summarized as follows:

- Computer vision: Introduced to computer vision fundamentals using OpenCV to detect faces, people, cars, and other objects, even when images are manipulated in rotations or scaling situations.
- Natural language processing: Explore the Natural Language Toolkit to process and extract text data: learning about the tokenization of words and sentences, part-of-speech tagging and stemming with lemmatization for the best analysis of textual data.
- Recommendation systems: Be introduced to various recommender systems, such as collaborative filtering with k-nearest neighbors, using either items or users.
- Networking and graphical models: Provide a systematic overview of methods for analyzing large networks, determining important structures in such networks, infer missing data in networks, and use graphical models to estimate and display a network of interactions.
- AutoML: Learn about Automatic Machine Learning (AutoML) methods and use PyCaret's regression module to predict continuous values using various techniques and algorithms.
- Big Data: Introduce Big Data and data engineering with the Hadoop ecosystem, the MapReduce paradigm, Apache Spark, and the up-and-coming Splunk.
- MLOps: MLOps, short for Machine Learning Operations, is an essential aspect of Machine Learning engineering that aims to teach the process of deploying, managing, and monitoring machine learning models in a production environment.

Opportunities and Challenges

The development of a data science curriculum for an engineering technology program comes with a number of opportunities and challenges to meet requirements and expectations.

Opportunities can briefly be summarized:

- Meeting the high demand for professionals: As the amount of data generated by businesses and other organizations continues to grow, there is a growing need for individuals with the skills to process, analyze, and make decisions based on this data.
- Integrating data science into existing engineering technology programs: By including data science as a critical component of engineering technology programs, institutions can help equip students with the latest skills and knowledge to succeed in the rapidly evolving tech landscape.
- Preparing students for the job market: Data science is a high-demand field with numerous job opportunities and competitive salaries. By completing a data science curriculum, students can position themselves to be competitive in the job market.

Challenges can briefly be summarized:

- Keeping up with a rapidly evolving field: Data science is a fast-moving field, and it can be challenging to stay up-to-date on the latest tools, techniques, and best practices.
- Ensuring student readiness: Many engineering technology programs are designed for students who have already completed coursework in computer science and mathematics. It can be challenging to develop a data science curriculum that is accessible to students who may have limited prior experience in these areas.
- Finding qualified faculty: As the demand for data science professionals continues to grow, the pool of qualified faculty in this field may be limited. Finding qualified instructors to teach data science courses can make it challenging.

Conclusion

Developing a data science curriculum for an engineering technology program is critical to preparing students for careers in this rapidly growing field. By integrating data science into the curriculum, engineering technology students will gain a comprehensive understanding of the concepts, methods, and tools necessary to work with data. This will enhance their technical skills and help them develop the ability to think critically and make data-driven decisions, which are essential for success in the data science field. Additionally, developing a data science curriculum for engineering technology programs will help meet the growing demand for a more educated workforce who can work in the data science field since data has to be collected, not just analyzed. With the proper training, engineering technology graduates will be well-positioned to tackle the complex data-related problems that organizations face today. By providing students with hands-on experience in data science techniques and tools, educational programs can help to produce the next generation of data-savvy engineers who will be capable of driving innovation and contributing to the advancement of their respective industries.

Acknowledgment

This work was supported in part by the Commonwealth Cyber Initiative (CCI), an investment in the advancement of cyber R&D, innovation, and workforce development in Virginia. For more information about CCI, visit cyberinitiative.org.

References

[1] Jared Dean. *Big data, data mining, and machine learning: value creation for business leaders and practitioners*. John Wiley & Sons, 2014.

[2] Yin Zhang, Dan Wu, Loni Hagen, Il-Yeol Song, Javed Mostafa, Sam Oh, Theresa Anderson, Chirag Shah, Bradley Wade Bishop, Frank Hopfgartner, et al. Data science curriculum in the ifield. *Journal of the Association for Information Science and Technology*, 2022.

[3] Sanjeevi Chitikeshi, Jake Hildebrant, Otilia Popescu, Orlando Ayala, and Vukica Jovanovic. Integrating Statistical Methods in Engineering Technology Courses. In *2018 ASEE Annual Conference Exposition*, 2018.

[4] Sebastian Raschka, Joshua Patterson, and Corey Nolet. Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information*, 11(4):193, 2020.

[5] Durga Prasad Penumuru, Muthuswamy Sreekumar, and Premkumar Karumbu. Identification and classification of materials using machine vision and machine learning in the context of industry 4.0. *Journal of Intelligent Manufacturing*, 31(5):1229–1241, 2020.

[6] Bureau of Labor Statistics. U.s. department of labor, occupational outlook handbook, data scientists. URL `https://www.bls.gov/ooh/math/data-scientists.htm`.

[7] Benjamin S. Baumer and Nicholas J. Horton. Data science transfer pathways from associate's to bachelor's programs. *Harvard Data Science Review*, 5(1):1229–1241, 2023.

[8] Malcolm Haynes, Joshua Groen, Eric Sturzinger, Danny Zhu, Justin Shafer, and Timothy McGee. Integrating data science into a general education information technology course: an approach to developing data savvy undergraduates. In *Proceedings of the 20th Annual SIG Conference on Information Technology Education*, pages 183–188, 2019.

[9] Richard D De Veaux, Mahesh Agarwal, Maia Averett, Benjamin S Baumer, Andrew Bray, Thomas C Bressoud, Lance Bryant, Lei Z Cheng, Amanda Francis, Robert Gould, et al. Curriculum guidelines for undergraduate programs in data science. *Annual Review of Statistics and Its Application*, 4:15–30, 2017.

[10] Vukica M Jovanovic, Murat Kuzlu, Otilia Popescu, Abdul Rahman Badawi, Deborah Kay Marshall, Salih Sarp, Spyridoula Tsouganatou, Petros J Katsioloudis, Linda Vahala, and Hongyi Wu. An initial look into the computer science and cybersecurity pathways project for career and technical education curricula. 2020.

[11] Murat Kuzlu and Otilia Popescu. Upgrading of a data communication and computer networks course in engineering technology program. In *2020 ASEE Virtual Annual Conference Content Access*, 2020.

[12] Otilia Popescu, Vukica Jovanovic, Sanjeevi Chitikeshi, and Isaac Flory. Introduction of Mechatronics Specialization through Concentration Areas in the Mechanical and Electrical Engineering Technology Programs. In *2018 ASEE Annual Conference Exposition*, 2018.

[13] Vukica Jovanovic, Alok Verma, and Mileta Tomovic. Developing an Option in Mechatronics System Design

within the Mechanical Engineering Technology Program. In *2013 International Forum on Systems and Mechatronics*, 2013.