IE 492 Projects – Spring 2022

Project 1: Prediction of KPIs using other indicators and setting the best values of the latter to optimize KPIs

N. Aras

(4 students - At least two students in the team must have taken IE 425/IE 582 and at least one student must have taken IE440.)

The project is based on predicting the value of any key performance indicator (KPI) using related activity levels/parameter collected from various sources and adjusting the values of these indicators so that the KPIs take on desired values. If the predicted value of the KPI meets the target set by the company, no action is required. However, if the targeted value of the KPI cannot be reached, then it will be necessary to determine at what level the activities to be carried out in order to reach the KPI. The first problem to be focused on in this project is the prediction problem, which is handled within the scope of data analytics/data mining. Note that each KPI is an output/response variable (dependent variable), and the realization levels of activities or the parameters are input variable/feature (independent variable). The aim is to find the function or relationship that estimates the value of the output variable with the smallest error by using the values of the input variables. For the solution of this problem, starting from the linear regression method in statistics, regression trees, random forests and gradient boosting machine (GBM) methods in the field of data mining/machine learning will be used.

If the estimated value of a KPI does not reach the targeted value, the second problem becomes the following: at which values of the input variables will the predicted value of the KPI reach the targeted value. Note that this is an optimization problem. If a function can be found that predicts the performance indicator value (e.g., with a linear/nonlinear regression model), then a nonlinear programming or integer nonlinear programming model can be used for solving the problem. However, it is known that linear regression method performs worse than machine learning methods such as regression trees, random forests and GBM in terms of prediction. If better prediction is obtained by such methods, non-derivative search methods such as Simplex Search has to be used to solve the optimization problem.

Data will be available within two weeks.

Project 2: Automation of route planning for last mile delivery

N. Aras, Ü. Bilge (3-4 students) - Assigned

Vivense sells furniture and home decoration products through online and offline channels. Company provides delivery and assembly services for some of these products to all cities in Turkey. This is managed by the delivery fleet in the warehouses. Efficient planning of the routes is very critical for customer satisfaction as well as cost optimization. Currently this planning process is managed by the route planners manually. This project is about automation of this process considering constraints in:

1) Type of service: Delivery, delivery & assembly, after-sales services (fixing product problems)
2) Location of customer
3) Total time needed (This is the typical bottleneck - especially for assembly)
4) Capacity of the car
5) On-time delivery
6) Cost of delivery (related to the distance travelled)

Project 3: Assessment of Alternative Course Registration Systems
G. Yücel

During the registration periods, students aim to register to a set of course from a common course pool in order complete their academic programs. Considering the finite capacity of the courses and diverse interests of the students, a protocol is needed in order to match students to courses (who is going to get which course?). In BÜ, the main matching protocol is first-come-first-served. Alternative matching protocols can also be used, such as bidding etc. The aim of this project is to evaluate the efficiency of alternative matching systems for course selection using agent-based simulation.

Project 4: Designing and implementing a decision-support system for determining material replenishment parameters
G. Yücel

Company X needs to manage the inventories of a large number of externally supplied components that are used in production. Considering the variability of consumption characteristics of these components, a one-size-fits-all replenishment policy cannot be used for all items. However, it is also not practical to analyze these components one by one to decide on the replenishment policy to be used. In this project, the team is expected to develop a decision support system that will use the historical consumption pattern of a component, analyze its characteristics, recommend a replenishment policy and its key parameters. The project will rely on basic data analysis, simulation and inventory control knowledge.

Project 5: Teller Quantity Optimization of a Banking System:
A. Korugan

A major bank operating in Turkey has to optimize the number of tellers working in a major bank branch. Arrival of customers fluctuates on a daily basis and within a day. There are multiple customer classes based on their value. The customer wait times cannot exceed a certain upper bound. Each customer class has a different tolerance for waiting. Also customer representatives try to sell products to customers while they are waiting for service. Although the bank wishes to minimize customer waiting times it does not intend to totally eliminate them.

Project 6: Forklift Quantity Optimization of a Float Glass Manufacturer:
A. Korugan

A float glass manufacturer uses forklifts as finished product buffers at the end of each float glass production line. Each line manufactures products in three different qualities, randomly. Since the production cannot be stopped when a buffer is full or not functional then the product related to this buffer will be discarded. Forklifts are charged regularly. They also fail and are repaired randomly. The objective is to minimize the absence time of forklifts as buffers by increasing the number of forklifts. Yet the cost of forklifts are significant.

Project 7: Scheduling traffic lights at Mecidiyeköy square
İ.K. Altınel (3-4 students)

This project aims to determine a green-red light schedule that optimize pedestrian and vehicular traffic flows at Mecidiyeköy square over a fixed planning horizon. Students will be responsible for the development and solution of an optimization model. The determination of model parameters will require the collection of data and its analysis. After its validation the model will be used to evaluate possible future scenarios.

Project 8: Minimum Spanning Tree Problem with Conflicts
İ.K. Altınel
(3-4 students, prerequisite: IE 456 or a basic background in graph algorithms)

A variant of the ordinary minimum spnanning tree problem, the Minimum Spanning Tree Problem with Conflict Constraints (MSTC) will be considered in this project. A conflict constraint states that a certain pair of edges cannot be contained simultaneously in a feasible solution. It is convenient to represent these conflict constraints in terms of the so-called conflict graph whose vertices correspond to the edges of the original graph, and whose edges represent conflict relations. Then, every stable set of the conflict graph is a conflict-free subset of edges. Hence, MSTC becomes the determination of a stable set of the conflict graph whose elements represent a connected acyclic and spanning subgraph of the original graph with minimum total weight. The goal of this project is to develop and implement Lagrangean heuristics for MSTC.

Project 9: Assignment Problem with Conflicts
İ.K. Altınel
(3-4 students, prerequisite: IE 456 or a basic background in graph algorithms)

A variant of the ordinary assignment problem, the Assignment Problem with Conflict Constraints (APC) will be considered in this project. A conflict constraint states that a certain pair of edges cannot be contained simultaneously in a feasible solution. It is convenient to represent these conflict constraints in terms of the so-called conflict graph whose vertices correspond to the edges of the original graph, and whose edges represent conflict relations. Then, every stable set of the conflict graph is a conflict-

free subset of edges. Hence, APC becomes the determination of a stable set of the conflict graph whose elements represent a perfect matching of the original bipartite graph with minimum or maximum total weight. The goal of this project is to develope and implement Lagrangean heuristics for APC.

Project 10: Modeling the dynamics of mucilage in Marmara Sea and potential solutions
Y. Barlas (3 students)

An alarming mucilage problem was observed in Marmara sea in late summer-early fall of 2021. But mucilage in Marmara sea (and other seas) is actually not a new problem. It has been a long-term 'dormant' dynamic problem that causes panic only when it surfaces. The purpose of the project is to model the underlying long-term (dormant) dynamics that generate mucilage. Once these biological and human-originated interactions are modeled and analyzed, potential long-term solution strategies will be investigated. The project would primarily utilize system dynamics modeling and policy analysis by simulation experiments.

Project 11: Interactive dynamic 'diving simulator' to train novice scuba divers
Y. Barlas and G. Yücel (3 students)

In scuba diving, the body is subject to several forces, some of which being non-linear and delayed. Thus, smooth scuba diving and stabilization is not trivial, which can be risky for the diver. The most essential factor in this process is the force exerted by the jacket (buoyancy compensator), so the diver regulates buoyancy by deflating air from or inflating air into the jacket. The basic forces and feedbacks involved in this process were already modeled in a M.S. thesis by Evrim Dalkıran (2006) and a simple interactive simulator was built. The purpose of this new project is to extend this prior work in two directions: i- to include and model other realistic factors and forces, such as the hand and fin movements of the diver, position and other diver characteristics, ii- to create a more realistic and richer scuba diving game, by using more advanced software, preferably with extensive web-based and animation features. Firstly the original model will be improved, then the interactive game will be built using this improved model, and finally the game will be thoroughly tested by players. The ultimate purpose is to develop a 'diving' simulator to help diving schools/clubs in training novice divers. The project will involve substantial amount of system dynamics simulation and general computer programming.

Project 12: Multi-Stage Inventory Planning for the Glassware Supply Chain
T. Bilgiç (3 students)

A glassware manufacturer produced to stock and owns a distribution center in Tuzla that feeds the market composed of large and small retailers. The task is to determine when and where to hold which inventory. You will work with real data at the SKU level and come up with hypotheses and models to propose a solution. Data handling and analysis, optimization or simulation model building and implementation skills are required.

Project 13: Exploring Vascular Networks using Graph Theory
T. Ekim (3 students)

Vascular networks are formed by blood vessels. The understanding of their structure is crucial for the diagnosis and treatment of several diseases including cancer. Accordingly, in this project, we aim to acquire a graph-theoretical understanding of the structure of vascular networks. In particular, we hypothesize that there are statistically meaningful differences between healthy and pathological tissues' vascular networks in terms of graph-theoretical parameters. This novel approach, proposed in Vilanova et al (2017) has not been fully explored yet although it has the potential to provide quantitative validation for the medical expertise and imply high clinical applicability.

The first step of this study consists in searching open-source biomedical imaging datasets (see below) to find appropriate data. Then the students will extract graph representations of vascular networks from these images using open-source software such as NEFI (http://nefi.mpi-inf.mpg.de), ImageJ (https://imagej.net/Welcome.) and Angiogenesis Analyzer.

Next, the students will conduct a detailed analysis of these extracted graphs by computing several graph-theoretical parameters such as the number of cycles per vertex, the average length of cycles, maximum/average degree, and density. More evolved parameters such as the treewidth, and the minimum feedback vertex set will also be studied.

Programming skills are required for this project. IE 456 recommended.

References:

Vilanova, G., Colominas, I. & Gomez, H. Computational Modeling of Tumor-Induced Angiogenesis. *Arch Computat Methods Eng* **24**, 1071–1102 (2017).
https://radiopaedia.org/articles/imaging-data-sets-artificial-intelligence
https://www.nature.com/articles/sdata2018180#Sec10
https://www.sciencedirect.com/science/article/abs/pii/S1477513120303193
https://www.sci.utah.edu/cibc-software/cibc-datasets.html
https://sites.google.com/site/hosseinrabbanikhorasgani/datasets-1
https://www.researchgate.net/post/Database_for_retinal_angiogram_images_and_which_database_is_better_for_selection

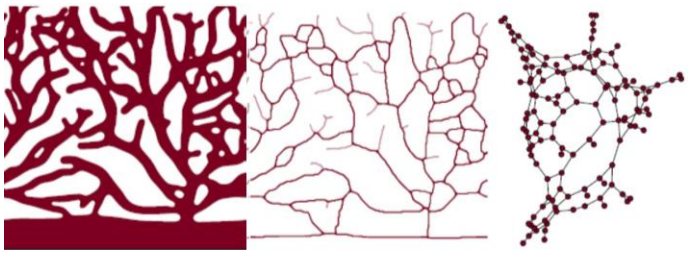

Fig. *(from Vilanova et al. (2017)) From a blood vessel image, the first stage called binarization transforms the image into a 0/1 matrix. The next phase called thinning (or skeletonization) makes the image of vessels thinner and usually all of the same widths. Then comes the graph extraction phase where we obtain a graph from the skeleton that represents the initial vascular network.*

Project 14: Personalized dynamic pricing in agricultural loan applications
M. G. Baydoğan
(3 students – At least one student in the team must have taken IE 425/IE 582)

Lenders offer different prices to different pricing segments in loan applications. This is due to the fact that different pricing segments have different variable costs. This cost difference is due to the customer risk. As a business objective, one of the main concerns is profit maximization in this context however there is a trade-off between pricing and customer loss under a competitive environment. In other words, competitor prices should also be taken into consideration. Agricultural loan applications have their unique characteristics due to the external factors such as seasonality, weather events and etc. Moreover, applicant's risk can be quantified based on the details of the loan requirement. For example, if a farmer needs loan to purchase a specific brand of fertilizer to produce wheat, you may quantify the risk based on the fertilizer characteristics, potential price after the harvest and suitability of the land (i.e. geographic and climatic conditions) and etc. In other words, loans are needed to make an investment (i.e. grow wheat) that will pay back after some period of time (i.e. 9 months) with associated risks. Hence there is a need for a personalized pricing specific to each farmer. This project is about learning price sensitivity of the applicants and pricing loans under these circumstances.

Project 15: Personalized product recommendations on second-hand platforms
M. G. Baydoğan
(3 students - All students in the team must have taken IE 425/IE 582)

With the advent of online marketplaces which millions of people worldwide visit and make purchase every second, the shopping experience and competition between these platforms have been significantly changed and recommendation systems have become a more critical part of these platforms and gained popularity in the literature. One of these online marketplaces, in which the recommendation system plays a key role, is second hand platforms. In addition to general recommendation problems, these platforms have several problems which are specific to this domain such as compromising extremely unique item sets that makes the problem difficult with respect to other domains. This project is about developing a product recommendation system using the clickstream data of users. Experience on dealing with huge amount of data is required.

Project 16: Physics informed machine learning for wind power forecasting tasks
M. G. Baydoğan and K. Bayramoğlu Kavlak
(3 students – At least one student (preferably two) in the team must have taken IE 350 and IE 425/IE 582 , familiarity with partial differential equations is suggested, experience in Pytorch is needed)

Wind power generation is always associated with uncertainties due to the fluctuations of wind speed. Accurate wind power forecasting is not only important for the efficient operation of power systems but also useful for power plant operators or participants making decisions in energy markets. Almost all the "good" forecasting approaches

make use of the weather forecasts as the input to the forecasting models. Weather forecasts are mainly obtained using numerical weather prediction (NWP) which consists of the set of mathematical models of the atmosphere and oceans to predict the weather based on current weather conditions. "NWP models employ a set of equations that describe the flow of fluids. These equations are translated into computer code and use governing equations, numerical methods, parameterizations of other physical processes and combined with initial and boundary conditions before being run over a domain (geographic area) [1]". Typical output of an NWP model is a combination of time, latitude, longitude, variable and level. A sample forecast can be as follows: the wind speed at 100m above ground at longitude of 28 and latitude of 37.25 at hour 8 of 2019-11-12. Here the wind speed is the variable, 100m above ground is the level. As NWP models are run globally with some approximations, they may not be providing reasonable forecasts for power prediction of a particular wind farm. On the other hand, historical power output of a wind farm is known and some adjustments on the provided forecasts can be performed using supervised learning methods. There is a vast literature on the use of machine learning for wind power forecasting tasks but they do not consider the inherent physics (i.e. associated equations to define the flow of fluids). The aim of this project is to develop wind power forecasting models which integrates the information from physics. For example, Navier–Stokes equations are a set of partial differential equations commonly used to characterize fluid dynamics in three-dimensional, unsteady form [2]. Using the physical facts, machine learning algorithms can be guided to obtain better/robust forecasts.

[1]https://www.weather.gov/media/ajk/brochures/NumericalWeatherPrediction.pdf accessed on 18.02.2022
[2] Constantin, P. and Foias, C., 1988. *Navier-stokes equations*. University of Chicago Press.

Project 17. PRODUCTIVITY AND SAFETY IMPROVEMENT OF A WORK SYSTEM
M. Ekşioğlu <Team of 3 to 4 Students>

Description: Application of ergonomics, lean six sigma and other industrial engineering principles to the work systems enhances productivity, quality, and safety and health. This project involves application of these principles in the evaluation and redesign of a selected work system. Project consists of three main parts: (1) data gathering for identification of productivity, safety and health issues, (2) solving the identified issues, and (3) cost-benefit analysis.

Requirement: IE 430 Ergonomics & Human Factors Engineering (all members)

Project 18. A HUMAN-COMUPTER INTERACTION DESIGN
M. Ekşioğlu <Team of 3 to 4 Students>

Description: For this project, students will choose an interface problem and go through a design process to brainstorm, conceptualize, develop, test, develop,

prototype, test, and report on your new or revised interface. Students will apply the design process, methods and principles learned in IE 48L course. You are to decide on a specific application within the selected domain, perform a literature search or background review on the topic, design an interface for the particular application chosen, and test and evaluate your design.
Project Domains: • Education • Medicine & Health Systems • Transportation • Home • Crime • Entertainment • Design • Manufacturing • Service • Banking • Business • Food Services • Agriculture • Sports • Other domains?
Some project ideas:

- Curated maps and routes (e.g. Spotify for walking or biking)
- Garden guide (e.g. planning, managing and tracking a garden)
- Back-end system for an animal shelter (for employees to manage animals)
- Back-end of restaurant menu ordering system (e.g. employee interface for adding inventory)
- Catalog store shopping assistant
- Recreation Centre app
- Ski hill application
- Mobile field guide for a park (e.g. for mushrooms, flowers, etc.)
- Driving data app
- Period tracking app
- Farmer market vendor app
- Something of your choosing

Requirement: IE 48L Human-Computer Interaction Design (all members)

Project 19: Understanding the dynamics of recovery from depression
H. Yaşarcan (3 students)

Assume a student who is going through a relatively difficult period as a result of several stress causing factors. This may cause the student to fall into depression. One may think that if the stress causing factors are once lowered to their critical levels that resulted in the depression, the student will immediately recover. However, that may not be the case as feeling depressed is an internal change of state in the student that must be managed and reversed. The group of students who will be assigned this project are expected to read the relevant literature about depression and construct a systemic feedback model that will be able to generate the dynamics of depression. If one understands the processes required to heal, it becomes easier to cope with the condition. Accordingly, the aim of the model that will be constructed in this project is to give an insight about the recovery dynamics of depression. If the recovery process is not understood by the student going through depression, the student may continue to keep his/her performance expectations unrealistically high. This will keep the stress levels high enough preventing the healing process to take place.

Projects assigned to some groups

1- Last mile warehouse location optimization

R. Güllü (4 students - Assigned)

Vivense delivers part of the furniture sold (especially big furniture like sofa, wardrobe, etc.) by its own delivery network. In this network, products are collected through suppliers into the main warehouse and then distributed to the last mile warehouses. Afterwards, products are delivered to customers by delivery teams from last mile warehouses. Vivense wants to revise its last mile warehouses network to achieve cost optimization and provide better and quicker service to the customers. This includes location identification and working model choice, own or franchise.

2- Decision Making Strategies for an agent trying to survive in the presence of a dominating rival in the market

H. Yaşarcan (4 students - Assigned)

Imagine a university dining hall that serves to students every midday. The dining hall has a fixed capacity and has a fixed daily service costs. It generates revenues that is calculated by multiplying the fixed meal price by the number of students having their lunch at the university dining hall. There is also a cafeteria that serves meals to students with a less waiting time in the queue and with a better quality. However, the capacity of the cafeteria is lower than the dining hall and its lunch prices are higher. The aim of this project is to study strategies for the cafeteria that will help it survive in the presence of its dominating rival that has a bigger capacity and bigger market share. A simulation model will be developed and it will be coded by R or Pyhton. Different strategies will be developed and their results will be obtained through simulations.