ADVANCES IN DATA SCIENCE & OPERATIONS RESEARCH

Virtual Conference

PRESENTED BY UNIVERSIDAD GALILEO IN COLLABORATION WITH INFORMSttt

From september 22 to 24

PROGRAM

Tuesday, September 22

09:00 - 9:30		Welcome
09:30 - 11:00	Workshop	Tableau & PowerBI -Self service en detalle (Julio Coronado & Marco Tulio Gómez)
11:00 - 11:30		Sponsor: Oportunidades en IEEE
11:30 - 12:00		Break
	Technical sessions	**Track in Data Science**
12:00 - 12:30		Data Science in Marketing: Practical Examples (Omar Martínez)
12:30 - 13:00		Bandidos Multibrazo Bayesianos y sus aplicaciones a Lead Gen (Carlos Zelada)
13:00 - 13:30		Transformers y BERT (Luis Leal)
13:30 - 18:00		Break
18:00 - 19:00	Keynote	Optimization with impact: my journey in public sector operations research (Laura Albert, Ph.D.)

Wednesday, September 23

09:15 - 9:30		Welcome
09:30 - 10:30	Keynote	Theory and Practice of Deep Learning (Boris Hanin, Ph.D.)
10:30 - 11:00		Sponsor: itm / Xumak
11:00 - 11:30		Break
11:30-13:00	Competition	Presentación de pósters
13:00-13:15		Premiación de la competencia de pósters
13:00-18:00		Break

		Track in Business Intelligence	**Track in Operations Research/Optimization**
18:00-18:30	Technical sessions	Exploración de datos con IBM Cognos Analytics (Héctor Asturias)	Enhancing the Wisdom of Crowds through Multimodal Input Elicitation and Principled Aggregation Models (Adolfo Escobedo, Ph.D.)
18:30-19:00		Cloud BI con Microsoft Azur y Power BI (Juan Manuel Alvarado)	Models and algorithms for fuel-constrained autonomous ground vehicle path planning problems under uncertainty (Saravanan Venkatachalam, Ph.D.)
19:00-19:30		MicroStrategy Hyperintelligence: Zero-click Analytics (Walter González)	A Simulation-Optimization Approach to Improve the Allocation of Airport Security Screening Resources in Airport Terminal Checkpoints (Eduardo Perez, Ph.D.)

Thursday, September 24

9:00 - 9:15		Welcome
		Track in Machine Learning
9:15-9:45	Technical sessions	Aplicando Machine Learning en tiempo real (Luis Fernando Valdeavellano)
9:45-10:15		De un análisis a un producto: cómo operacionalizar mis modelos usando procesamiento en la nube (Rodrigo Rivera Ávila)
10:15-10:45		Productivización: de Idea a Modelo AI a producto para el cliente (Nery Guzmán)
10:45-11:15		Lanzamiento Capítulo Estudiantil de Univeridad Galileo-INFORMS
11:15 12:00		Break
12:00 - 13:30	Workshop	Introducción a Data Science con Julia (Preng Biba)
13:30-18:00		Break
18:00-19:00	Keynote	Digitalización de la Ciencia y la tecnología en Costa Rica: retos de la política pública para una toma de decisiones basada en datos (Federico Torre-Carballo, Ph.D.)
19:00-19:30		Close event

Keynote speakers

Laura Albert, Ph.D.

Optimization with impact: my journey in public sector operations research

Language: English **09/22/2020** **18:00 Guatemala (UTC-6)**

Government programs spanning public safety, transportation security, and critical infrastructure protection must deliver essential services by managing risks such as health emergencies, crime, acts of terrorism, and natural disasters. Doing so requires allocating resources in complex systems that span people, processes, vehicles, and critical infrastructure, where many decisions are interrelated.

Government leaders and researchers have been studying how to design and operate public sector systems to manage risk for the last half a century. Although researchers have created a body of knowledge for supporting prescriptive and predictive decisions in the public sector, public safety leaders must continually adapt to address new risks in budget-constrained environments. As a result, many research challenges remain. In particular, the SARS-CoV-2 pandemic has provided many challenges for public safety leaders in terms of providing services, reducing risk to first responders and to the public, and addressing new challenges that have emerged due to the pandemic.

In this talk, Dr. Laura Albert will discuss her research that studies how to design and operate public sector systems using optimization methodologies. She will discuss how she has connected theory and modeling to application in applications in the United States ranging from emergency medical services, aviation security, and critical infrastructure protection. She will also discuss policy insights as well as insights obtained from putting the results into practice in real world settings.

.

Boris Hanin, Ph.D.

Theory and Practice of Deep Learning

Language: English **09/23/2020** **9:30 Guatemala (UTC-6)**

In the past decade machine learning models based on neural networks have achieved state of the art performance on a variety of machine learning tasks, ranging from computer vision (e.g. self-driving cars) to natural language processing (e.g. Siri, Google Translate, etc), and reinforcement learning (e.g. AlphaGo). In this talk I will survey these successes as well as some outstanding empirical and theoretical challenges.

Federico Torres-Carballo, Ph.D.

Digitization of Science and Technology in Costa Rica: Public Policy Challenges for Data-Based Decision Making

Language: Spanish **09/24/2020** **18:00 Guatemala (UTC-6)**

A national ecosystem of information and communication technologies goes through several stages to consolidate the benefits of a network economy. This conference will address the ICT ecosystem of Costa Rica in the field of science, technology and innovation (STI), its initial evolution from disjointed steps in different institutions in the sector, the disruptive effect of OECD commitments in terms of digitization and data services in STI, the emergence of integration initiatives such as the National Science and Technology System (SINCYT), its link to national public policy and the challenges that a data-based decision-making model demands.

Track 1: Data Science

Lic. Omar Martínez

Data Science in Marketing: Practical Examples

Language: Spanish **09/22/2020** **12:00 Guatemala (UTC-6)**

In this technical session, the unconventional use of data science applied to marketing will be presented. In particular, an overview will be presented of how Data Science, Machine Learning and Artificial Intelligence techniques and methods can be integrated into analysis and problem solving in marketing. In addition, the fundamentals and principles of interpretability and reproducibility in statistical learning models will be studied in depth. Some topics to be covered are: the difference between marketing analysis, machine learning and statistical learning. The key attributes of a data scientist specializing in marketing such as: extracting valuable information from any data set, including a simple export of Google Analytics and Google Search Console and some resources to start your professional development if you want to dedicate yourself to this area.

Track 1: Data Science

Ing. Carlos Zelada

Bayesian Multi-Arm Bandits and their applications to Lead Gen

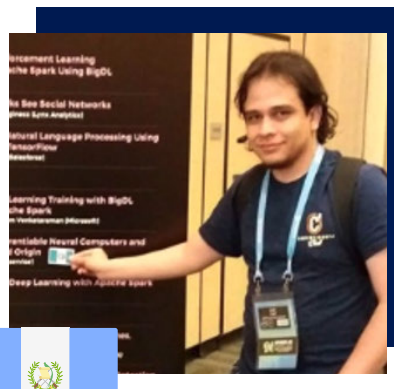
Language: Spanish **09/22/2020** **12:30 Guatemala (UTC-6)**

There are many options for ad providers in the lead gen industry and they in turn can display a large number of ads. Deciding which ad or which ad provider to display can be complicated, as you may have very little information about your conversion rate. Likewise, it is not possible to use only the historical data because every day the demand changes and you have to adjust to it. The multi-arm bandit method, named after a casino slot machine, can help solve this problem. The broad idea is that by not knowing the reason for paying a machine or a set of them, so when you start playing in all of them and as payments are obtained, more weight is given to those that are paying. If you keep this strategy for a while you will find the machine that pays the most. In this technical session, we will talk about bandits, bayes and marketing.

Ing. Luis Leal

Transformers & BERT

Language: Spanish **09/23/2020** **13:00 Guatemala (UTC-6)**



For a long time recurrent neural networks (and later the use of attention mechanisms in them) were the state of the art for sequential modeling, including Natural Language Processing (NLP) and language models. Recently a new type of model changed the world of NLP and sequential modeling, creating a new family of algorithms and models based on one that gave rise to all: Transformers and, in turn, its most successful variation BERT (bidirectional encoder representation from transformers) has served as the basis for recent advances in artificial intelligence. In this talk will address a bit of Transformers, their fundamental ideas and how they completely modify and renew sequential modeling using natural language processing cases as examples.

Ing. Preng Biba

Introduction to Data Science with Julia

Language: Spanish **09/24/2020** **12:00 Guatemala (UTC-6)**

Julia is one of the most recognized languages currently by the scientific community to develop applications where exhaustive numerical computation is required. The main reason for this boom lies in the combination of four aspects for the development of scientific applications that Julia manages to combine, these are: the ease and dynamism of Python together with the speed and efficiency of C. For this reason, Julia has become in a trend to solve large-scale applications in data science and machine learning. In this workshop, an introduction to its syntax will be presented and some of its most relevant characteristics for the implementation of solutions within the aforementioned areas will be evaluated.

Track 2: Operations Research / Optimization

Adolfo R. Escobedo, Ph.D.

Enhancing the Wisdom of Crowds through Multimodal Input Elicitation and Principled Aggregation Models

Language: English **09/23/2020** **18:00 Guatemala (UTC-6)**

The tendency of groups to outperform individuals is a principle commonly referred to as the “wisdom of crowds”. Human computation is an evolving related research area that studies how to harness human intelligence to solve difficult problems. Because of varying subjective scales among humans, more reliable outcomes can be obtained by aggregating inputs of many individuals into collective estimates. This talk discusses how the joint aggregation of cardinal inputs (i.e., numerical estimates) and ordinal inputs (i.e., ranking estimates) can be used to further enhance the effect of the wisdom of crowds in human computation. We demonstrate how cardinal inputs can increase the quality of collective ranking estimation and how multimodal information (i.e., cardinal and ordinal inputs) can increase the quality of collective numerical estimation. Moreover, we leverage these multiple modalities to extract wisdom from smaller crowds to make human computation more practical for companies to implement. Lastly, we demonstrate the benefits of consensus-based optimization models relative to these goals.

Track 2: Operations Research / Optimization

Saravanan Venkatachalam, Ph.D.

Models and algorithms for fuel-constrained autonomous ground vehicle path planning problems under uncertainty

Language: English **09/23/2020** **18:30 Guatemala (UTC-6)**

Use of heterogeneous unmanned ground vehicles (UGVs) is increasing in both civil and military applications. This talk presents two variants for a fuel-constrained vehicle path-planning problem in the presence of multiple refueling stations: a deterministic model to minimize the maximum utilization of any UGV; and a two-stage stochastic model under uncertainty in availability of UGVs. Given a set of points of interests (POI), a set of refueling stations for UGVs, and a depot where the UGVs are stationed and their availability is random, the objective is to determine route for each UGV starting and terminating at the depot such that overall incentives collected by visiting POIs is maximized or the travel distance is minimized. We use an outer approximation based decomposition algorithm and a genetic algorithm based heuristic to solve large instances. Along with extensive computational results, a data driven simulation study is performed using robot operating system (ROS) framework to corroborate the use of the proposed mathematical models.

Eduardo Pérez, Ph.D.

A Simulation-Optimization Approach to Improve the Allocation of Airport Security Screening Resources in Airport Terminal Checkpoints

Language: English **09/23/2020** **19:00 Guatemala (UTC-6)**

In this research, a simulation-optimization strategy is developed to improve the operation of airport checkpoints. Simulation-optimization is a suitable framework for problems involving data uncertainties that evolve over time, requiring important system decisions to be made prior to observing the entire data stream. This is indeed the case in airport security checkpoints, where the passenger arrival times are difficult to predict and requirements for equipment and human resources must be scheduled in advance. The team explicitly considered the uncertainties associated with future passenger arrivals and availability and performance levels of the resources in computing staffing and system configuration decisions. A computational study is presented that shows how a simulation-optimization strategy can be useful in planning resource allocation in airport checkpoints. The simulation-optimization strategy provided a 31.4% improvement for the passengers' cycle time when compared to a benchmark scenario. The computational study also reports the impact of adding resources to the checkpoint and the timing of adding those resources. The results of this research are expected to help improve the operation and safety of airport facilities by developing new decision-making models that will leverage technology and knowledge to enhance the operational effectiveness of airport security checkpoints. Equipping these facilities with better planning tools will allow for better-informed downstream distribution decisions. The resulting models and methods from this research will provide valuable information that will increase the potential of airport facilities to meet the organizational objectives under multiple operational scenarios.

Track 3: Business Intelligence

Lic. Julio Coronado y Lic. Marco Tulio Gómez

Tableau & Power BI the Self-Service in detail

Language: Spanish **09/22/2020** **09:30 Guatemala (UTC-6)**

Self Service is one of the strategies most used today in business intelligence architectures. In this workshop, you will start by defining the self-service strategy, its benefits and the challenges it represents. Then, it will lead to a review of the two products with the most market share: Tableau and Power BI. The goal is to understand how the self-service strategy can be executed effectively and independently of the product.

Lic. Héctor Asturias

Exploring data with IBM Cognos Analytics

Language: Spanish **09/23/2020** **18:00 Guatemala (UTC-6)**

This talk will explore IBM Cognos Analytics, this is a web-based business intelligence platform aligned by artificial intelligence that supports the entire analytics cycle.

Track 3: Business Intelligence

Lic. Juan Manuel Alvarado
Cloud BI with Microsoft Azure & Power BI

Language: Spanish **09/23/2020** **18:30 Guatemala (UTC-6)**

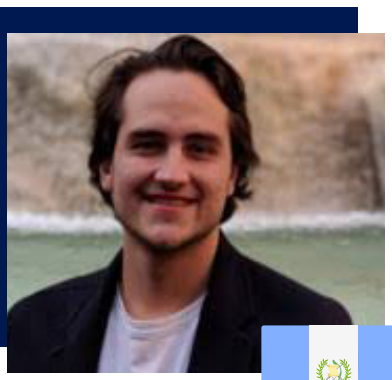
Currently the cloud is one of the most used resources in IT strategies due to its great technical and economic benefits. For this reason, different possibilities have been developed to implement a business intelligence solution in the cloud. All providers offer various products for this purpose. In this session, you will explore Microsoft's offering in relation to its Azure services such as SQL Database, Data Factory, Analysis Services, and Synapse.

Ing. Walter González
MicroStrategy Hyperintelligence: Zero-click Analytics

Language: Spanish **09/23/2020** **19:00 Guatemala (UTC-6)**

MicroStrategy has been known for being an independent provider in the world of business intelligence solutions. They have recently launched their hyperintelligence strategy where MicroStrategy allows integrating the world of information analysis between different providers. In this talk, the most relevant details of said strategy and its various functionalities will be presented.

Track 4: Machine Learning



Ing. Luis Fernando Valdeavellano
Applying Machine Learning in Real Time

Language: Spanish **09/24/2020** **09:15 Guatemala (UTC-6)**

In this talk, different tools and strategies used to evaluate machine learning models on streaming data and the challenges that this represents will be presented. In particular, connection to data sources, batch mode training, resource management and monitoring will be discussed. It is worth mentioning that the subject will be addressed with examples of real implementations that have been developed in the telecommunications industry in different Latin American countries.

Ing. Rodrigo Rivera Ávila
From an Analysis to a product: How to Operationalize My Models Using Cloud Processing

Language: Spanish **09/24/2020** **09:45 Guatemala (UTC-6)**

One of the main challenges we encounter as data scientists is presenting our own findings. We are experts at seeing what no one else sees, but we have difficulty replicating that knowledge across organizations. In this talk we will explore how to use Shinyapps and Amazon Web Services (AWS) to move from generating static data analytics to data products that generate value for the organization.

Nery Fernando Guzmán
Productivization: From an idea, to an Algorithm, to a Product Ready for the Market

Language: Spanish **09/24/2020** **10:15 Guatemala (UTC-6)**

The objective of the talk is to address the experience of managing a 100% Guatemalan team developing a SAAS (Software as a Service) platform that operates with 3 algorithms, two optimization algorithms and a prediction algorithm. This talk will discuss the validation of academic methodologies for DevOps, as well as the most important considerations and strategies to develop a product based on data insights; highlighting the lessons learned throughout this process.

Poster Session

Speakers
Rubén Darío González Monterroso
Edwin Estuardo Zapeta Gómez
Juan Pablo Carranza Hurtado
Sergio José Barrios
Diego Fernando Valle
Elda Magally Calderón
Lilian Carrera
Jose Fernando Pérez Pérez
Jonathan de León

Days:

September 22 to 24

Place:

Virtual conference, ZOOM WEBINARS

In collaboration with:

Sponsors: