DLI Accelerated Data Science Teaching Kit

Lecture 1.4 - Learning Goals and Expectations

Learning Goals

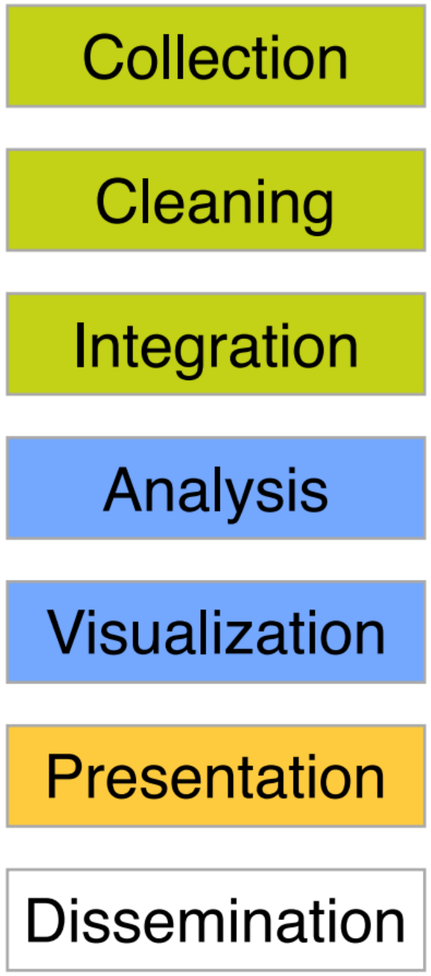
- Learn **visual** and **computation** techniques and use them in **complementary** ways
- Gain a **breath** of knowledge
- Learn **practical** know-how by working on **real data & problems**

Topics Covered

(Analytics Building Blocks)

- Collection
- Cleaning
- Integration
- Analysis
- Visualization
- Presentation
- Dissemination

Building Blocks. NOT Rigid Steps.



Can skip some

Can go back (two-way street)

- Data types inform visualization design
- Data informs choice of algorithms
- Visualization informs data cleaning (dirty data)
- Visualization informs algorithm design (user finds that results don't make sense)

“Caveats” for Students

- Learning data science can involve **a lot of programming**
- You need to be prepared to **learn many things** in short amount of time, e.g., visualization, JavaScript, CSS, Scala
- **Very common in industry**

From Previous Offerings …

The course materials of *CSE 6242 Data and Visual Analytics* at Georgia Tech are based on this teaching kit

- Class projects had turned into research publications, e.g., KDD, IUI
- Projects as portfolio pieces on CV
- Increased job and internship opportunities
- Receive numerous “thank you” notes from former students

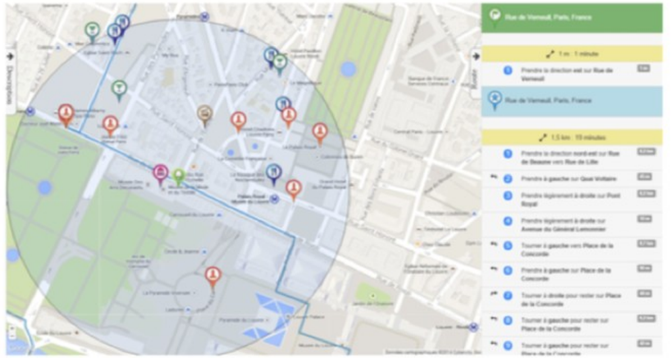
Aurigo: An Interactive Tour Planner for Personalized Itineraries

Alexandre Yahi; Antoine Chassang; Louis Raynaud; Hugo Duthil; Duen Horng (Polo) Chau
Georgia Institute of Technology
{alexandre.yahi, antoine.chassang, l.raynaud, hduthil, polo}@gatech.edu

ABSTRACT
Planning personalized tour itineraries is a complex and challenging task for both humans and computers. Doing it manually is time-consuming; approaching it as an optimization problem is computationally NP hard. We present Aurigo, a tour planning system combining a recommendation algorithm with interactive visualization to create personalized itineraries. This hybrid approach enables Aurigo to take into account both quantitative and qualitative preferences of the user. We conducted a within-subject study with 10 participants, which demonstrated that Aurigo helped them find points of interest quickly. Most participants chose Aurigo over Google Maps as their preferred tools to create personalized itineraries. Aurigo may be integrated into review websites or social networks, to leverage their databases of reviews and ratings and provide better itinerary recommendations.

Author Keywords
User Interfaces; Visualization; Recommendation; Tour itinerary planning

ACM Classification Keywords
(e.g. HCI): User interfaces



IUI'15 Full conference paper

ISPARK: Interactive Visual Analytics for Fire Incidents and Station Placement

Subhajit Das, Andrea McCarter, Joe Minieri, Nandita Damaraju, Sriram Padmanabhan, Duen Horng (Polo) Chau
Georgia Tech
Atlanta, GA, USA
{das, andream, jminieri, nandita, sriramp, polo}@gatech.edu

ABSTRACT

In support of helping to reduce the response time of firefighters, and thus deaths, injuries, and property loss due to fires, we introduce ISPARK. The ISPARK system determines where fire stations should be located, analyzes the primary causes of fires, the existing infrastructure, and response times, by using visualizations which show the GIS mapping of fire stations on a dashboard. Incidents and response times are shown as additional layers, with clustering of fire incidents to determine predicted fire station locations, forecasting of fire incidents using regression, causal, infrastructure, and personnel analysis, creating an interactive, multi-faceted method for locating fire stations. A comparison of urban and rural fire incident response times is another dimension of this study. We demonstrate ISPARK's usage and benefits using a publicly available dataset describing 300,000 fire incidents in the states of Massachusetts and Maine. ISPARK is generalizable to other geographic areas

Figure 1: Screenshot of ISPARK showing actual (pink) and predicted (green) fire station locations n Maine determined by our approach, using coordiates with actual driving distances from fire stations o actual fire incidents. Fire incidents are shown as small yellow dots. ISPARK reduces the average

KDD'15 Workshop paper

PASSAGE: A Travel Safety Assistant With Safe Path Recommendations For Pedestrians

Matthew Garvey
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mgarvey6@gatech.edu

Meghna Natraj
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
mnatraj@gatech.edu

Nilaksh Das
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
nilakshdas@gatech.edu

Bhanu Verma
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332, USA
bhanuverma@gatech.edu

Jiaxing Su
College of Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA
Jiaxingsu@gatech.edu

Abstract
Atlanta has consistently ranked as one of the most dangerous cities in America with over 2.5 million crime events recorded within the past six years. People who commute by walking are highly susceptible to crime here. To address this proble… …roup has developed a mobile application, PASSAG… …ed crime data to find "safe pat… Atlanta. user int…

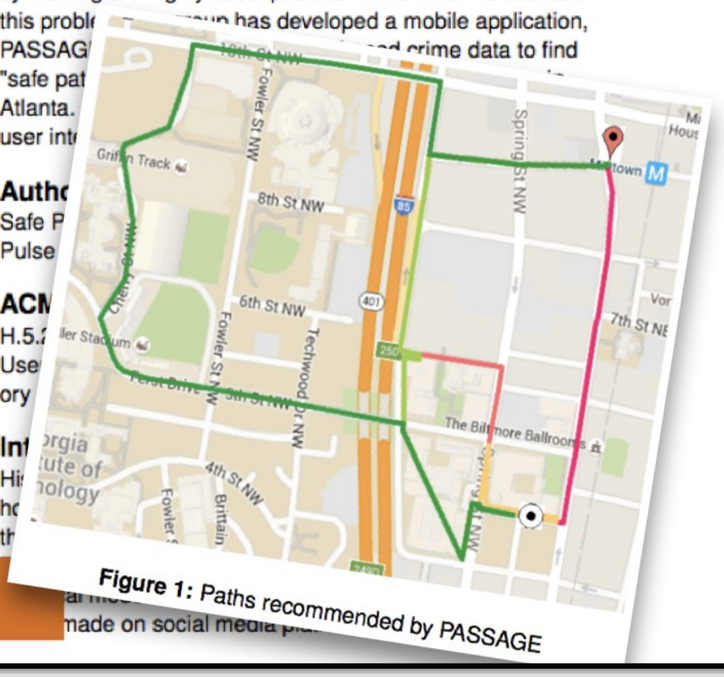


Figure 1: Paths recommended by PASSAGE

IUI'16 Poster paper

NVIDIA DEEP LEARNING INSTITUTE | Georgia Tech | Prairie View A&M University

Student Reflections

*"I feel like the concepts from your class are like a **rite of passage for an aspiring data scientist**. Assignments lead to a feelings of accomplishment and truly progressing in my area of passion."*

*"I really get more intuition about how to **deal with data with some powerful tools in HW3** [uses AWS]. That feeling is beyond description for me."*

*"I would like to say thank you for your class! Thanks to the skills I got from the class and the project, **I got the offer**."*

DLI Accelerated Data Science Teaching Kit

Thank You