

Proceedings of the Symposium for Virginia
Tech Undergraduate Research in Computer
Science

VTURCS 2019

April 30, 2019

6:30pm-8:30pm

Owens Banquet Hall

The Virginia Tech Undergraduate Research in Computer Science (VTURCS) program highlights the research and capstone projects of our undergraduate students during the past year. We are grateful for the support of Virginia Tech's Computer Science Resources Consortium (CSRC), a collection of companies that work with the Department of Computer Science to further our teaching and research mission.

R1: PairWise: Providing Accountability for Distributed Work	5
R2: A Case for Node.js Thread Pool Improvement	6
R3: GroundTruth: Enabling Experts to Geolocate Images by Working With Novice Crowds	7
R4: Rating and Rankings of College Basketball Teams	8
R5: Ascertaining the Accuracy/Cost Tradeoffs of Exploiting Functional Equivalence for Image Processing	9
R6: CORBA Slayer: From Centralized to Distributed Without Programming	10
R7: Augmenting OpenDSA eTextbook with Programmed Instruction Frames to enhance students interaction and engagement	11
R8: CrowdIA: Supporting Intelligence Analysis with Crowdsourcing and Context Slices	12
R9: Promoting Tolerance & Civility in Online Discourse through Re-Designing Button Labels	13
R10: Virginia Forest Visualization	14
R11: OpenDSA eTextbook to Enhance Student Understanding of B+ Tree	15
R12: Algorithmic Composition in Virtual Reality_WEI	16
C1: Consensus Clustering of Connectomes	17
C2: Multi-Feature Analysis of Research Papers to Classify Retracted Studies	18
C3: Tourism Destination Websites	19
C4: An Analysis of the Sexist and Racist content in Gaming Channels by Genre	20
C5: Louisiana French	21
C6: Street Art: an immersive VR experience	22
C7: Expanding HEART using Advanced Advertising Techniques	23
C8: Picturing the Trails	24
C9: BarkLight: Ubiquitous Computing in a Naturalistic Environment	25
C10: UI/UX Improvements to MatWeb	26
C11: Healthy Hokies	27
C12: Tree Viz	28

C13: BirdFeed	29
C14: Adult Day Services Memory Masterclass Promotional Video	30
C15: CSRC Interview Check-in System	31
C16: Fostering Digital Cooperative Management on the Appalachian Trail	32
C17: Department of Computer Science Absence Survey System	33
C18: Functional Subgraphs in Brain Networks	34
C19: Sustainability Values: Interactive Website for Values Diagnostic Reporting and Analysis	35
C20: Generative Mathematical Models for Real World Connectomes	36
C21: Accessibility Map @ VT	37
C22: PrepLab	38
C23: Cranial Genetics: Investigating Clustering Algorithms to Decode Gene Expression Data from the Human Brain	39
C24: Navigation Routing in Brain Networks	40
C25: The Impact of Reddit's Coverage of the 2018 Parkland School Shooting on the Political Polarization of Online Communities	41
C26: Online Role-related User Classification on Twitter	42
C27: The Effect of News Corrections in Today's Online Information Sharing	43
C28: PolitiGauge	44
C29: Influence Analysis of Information Operations on Twitter	45
C30: Analyzing the Language of Anti-Vaccination Pages	46
C31: Python4ML	47
C32: The Mayfair Project	48
C33: Adult Day Services Video Assessment	49
C34: Engineering a Cardboard Database	50
C35: Information and Computer Technology for Indigenous Knowledge Healthcare	51
C36: Gamifying Good Behavior	52
C37: Workflow For Developing Gestures for Virtual Reality Using Machine Learning	53
C38: Tobacco Testimonies	54

C39: Rentian Scaling in Brain Networks

55

C40: Trail Tweets

55

R1: PairWise: Providing Accountability for Distributed Work

Authors: Christopher Hurt (chris526@vt.edu)

Advisors: Jacob Thebault-Spieker (jthebaultspieker@vt.edu), Kurt Luther (kluther@vt.edu)

Abstract: Distributed online work is becoming increasingly impactful in today's information age. Large social media companies like Facebook, YouTube, and Reddit rely on distributed content moderation to maintain a healthy and fair user environment. These mechanisms, however, risk being flawed by the social behaviors of their human participants, particularly due to their political ideologies. This can lead to political biases in content moderation. This study analyzes this effect and potential solutions using the PairWise platform, specifically in the context of social media posts.

R2: A Case for Node.js Thread Pool Improvement

Authors: Jonathan Alexander jma@vt.edu

Advisors: Dongyoon Lee dongyoon@cs.vt.edu James Davis davisjam@vt.edu

Abstract: This project aims to investigate the potential for improvements to Node.js performance by altering thread pool utilization. Currently, Node.js routes all blocking IO tasks and CPU intensive operations to either the V8 or LibUV thread pool based on the origin of the task. By splitting these tasks between the thread pools based on task type, task origin, or some other metric there could potentially be improvements to the performance of Node.js applications. To determine if changes to the Node.js thread pool would yield improvements in performance we have compiled multiple benchmarks simulating real-world Node.js use cases. These benchmarks have been run with a prototype Node.js binary that allowed for the monitoring and manipulation of thread pool usage. Using this binary, we were able to monitor the thread pool profile of these benchmarks and alter the way tasks were partitioned between thread pools. Our data shows that there are multiple types of tasks with significantly different run durations and operations being utilized in real-world Node.js applications. Furthermore, data from the benchmarks showed that changing the way in which tasks are routed to a thread pool can impact the performance of the application. From this, we conclude that changes to the Node.js thread pool can impact application performance in a positive and significant way and that further research into what changes will improve performance could yield viable results.

R3: GroundTruth: Enabling Experts to Geolocate Images by Working With Novice Crowds

Authors: Daniel Ocheltree djo96@vt.edu Puriwat Lahpong pulahpon@vt.edu

Advisors: Kurt Luther kluther@vt.edu

Abstract: Photo and video verification is a key step in investigations in fields like journalism and human rights advocacy. Geolocation is a subset of image verification, where expert investigators identify the exact geographic location where a photo or video was taken. However, given time-sensitive information to break the news front and a minimal number of experts, this manual process doesn't scale easily and experts are often overwhelmed by the amount of information they must process -- not to mention the misinformation amongst the data. GroundTruth was created to help decrease the amount of satellite imagery that experts need to search through with the use of crowd workers, which we have chosen due to its large-scale and fast-speed nature coupled with a simple recruitment process. Throughout this semester, we revamped both the front-end and back-end of GroundTruth to enable a public release of the tool to journalists and other investigators. We wanted to create a consistent look and brand, as well as a more straightforward tool. This led to front-end work that included a visual overhaul, new logo, and improved user experience. We also wanted to increase the security of the webpage and allow for user's feedback. This leads to back-end work that included hashing IDs, modifying the database, and shifting to cloud storage on AWS. In future work, we will publically launch the tool and have real-world experts launch investigations. We will also have instrumentation set up to track how experts use GroundTruth so we can better understand what the needs of our users are and how we can cater to them. Finally, we will further look into improving the security of the webpage and application, ensuring a safe environment to work in.

R4: Rating and Rankings of College Basketball Teams

Authors: Lynda Nguyen (lnguyn8@vt.edu)

Advisors: Lizette Zietsman (lzietsma@vt.edu) Chris North (north@cs.vt.edu)

Abstract: Historically, sports and statistics goes hand-in-hand for game and player analysis, but we want to go farther than just the stats sheet. With such data, set algorithms and methods can attempt to predict the outcome of a game or entire series. Analyzing and managing data is crucial in this process, therefore, our research looks to optimize and compare certain attributes of the algorithms we looked at. The research conducted will compare various approaches to the rating and ranking of collegiate basketball teams. The systems of rankings will encompass the competitiveness and predictability surrounding March Madness. It will include methods such as Massey's method, Pomeroy's method, Logistical Regression and Markov Chain (LRMC), Colley's Method and in addition to various machine learning techniques. With statistical analysis through machine learning one will uncover the importance of features and learn their relationships. Machine learning will also determine how a team's past statistics can help determine predictability when facing new opponents. When looking at the environment that March Madness yields, there are multiple differences that the statistics of a team's season cannot account for. This study will allow one to investigate the importance of attributes versus models and raises the question if there is an upper limit to predictive quality.

R5: Ascertaining the Accuracy/Cost Tradeoffs of Exploiting Functional Equivalence for Image Processing

Authors: Shuangyi Li, amnos@vt.edu

Advisors: Aabhas Bhatia aabhasb@vt.edu, Zheng Song songz@vt.edu, Eli Tilevich tilevich@vt.edu

Abstract: If different services can provide the same application functionality, they are equivalent and present implementation choices to the developer. This research is concerned with empowering the developer with empirical findings to be able to answer the following question: Given the QoS requirements and a set of equivalent services, which combined execution of these services would best satisfy these requirements? To detect the presence of a face in an image, both cloud services provided by different vendors and edge services based on open-source projects implement this functionality. We conducted an experimental study of how the combined execution of equivalent services affects the QoS of real-world application scenarios. In particular, we analyzed the accuracy of face detection by combining the executions of three widely-used web services and two open-source implementations. To these executions, we applied the majority voting, weight voting, and stacking accuracy enhancement strategies. Based on the observed results, we defined practical guidelines to help developers identify the best execution strategy that satisfies given QoS requirements.

R6: CORBA Slayer: From Centralized to Distributed Without Programming

Authors: Kibur Girum, kibur1@vt.edu Graham Wobig, grahamwo@vt.edu

Advisors: Breno Dantas Cruz, bdantasc@cs.vt.edu Yin Liu, yinliu@cs.vt.edu Dr. Eli Tilevich, tilevich@cs.vt.edu

Abstract: Distributed processing in resource-constrained systems requires lightweight middleware, while mainstream middleware platforms (e.g., CORBA) impose a prohibitively high overhead to be integrated with such systems. To satisfy the stringent performance and resource usage requirements of resource-constrained systems written in C/C++, this project designs, develops, and evaluates a lightweight middleware framework for distributed processing. Our middleware framework--CORBA Slayer--features a parameterizable code generator with three essential features: a) generation of lightweight distributed functionality for two processes to communicate with each other through shared memory; b) powerful meta-programming facilities, exposed to the programmer by means of C++ annotations; c) customized marshaling into memory buffers for standard C-language types, including structs and simple dynamically allocated data structures. CORBA Slayer both reduces programmer effort while enabling efficient and robust distributed interaction for resource-constrained systems. We have implemented CORBA Slayer by leveraging LLVM to generate code for distributed embedded systems. Our approach is unique in automatically generating and transforming code to introduce distribution to centralized systems. Any developer facing the challenges of developing distributed systems for resource-constrained environments would benefit from using CORBA Slayer, as it provides programmer-friendly yet powerful automated distribution.

R7: Augmenting OpenDSA eTextbook with Programmed Instruction Frames to enhance students interaction and engagement

Authors: Samnyeong Heo, hsn1017@vt.edu

Advisors: Cliff Shaffer, shaffer@vt.edu Mostafa Mohammed, profmdn@vt.edu

Abstract: OpenDSA provides infrastructure and materials to support courses in a wide variety of Computer Science-related topics. Traditionally, OpenDSA uses slideshows to explain and visualize content such as Data Structures and Algorithms, Formal Languages, Finite Automata, and Programming Languages. OpenDSA allows students demonstrate knowledge of an algorithm by completing interactive exercises. However, static presentation of abstract, math-heavy material can be easily abused or skipped by students who quickly click through or skip the slides, or otherwise ignore the content. In general, static presentation is not an effective way for students to learn demanding materials such as is presented in a Formal Languages course. We propose to address these problems through use of Programmed Instruction (PI) pedagogy. This approach provides small segments of content (maybe a sentence or paragraph) frequently interspersed with questions or exercises. Therefore, students must actively engage with the slides by reading the contents and fully understanding the information in order to move forward by answering the question. We implemented a frame-based system to support PI pedagogy, making the slideshows far more interactive.

R8: CrowdIA: Supporting Intelligence Analysis with Crowdsourcing and Context Slices

Authors: Chandler Manns (chandm8@vt.edu)

Advisors: Tianyi Li (tianyili@vt.edu) Kurt Luther (kluther@vt.edu) Chris North (north@vt.edu)

Abstract: Making sense of large text datasets is a difficult problem in many domains that does not scale well for individuals. Crowdsourcing presents new opportunities for large-scale sensemaking. Most crowdsourced sensemaking approaches or NLP models focus on solving local tasks that are part of bigger problems. They also require expert intervention because of errors and bottlenecks that would otherwise harm the output quality. Mitigating these errors and bottlenecks would significantly reduce the burden on experts and increase the scalability of the sensemaking process, yet little is known about the types of mistakes, their contributing factors, and how they propagate to later analysis and impact the overall outcome. With CrowdIA, we build a pipeline that supports transient, novice crowd workers to contribute to a holistic sensemaking process asynchronously and meaningfully. Through this system, we analyze the mistakes and classify types of crowd errors to show how the amount and quality of input data influence worker performance.

R9: Promoting Tolerance & Civility in Online Discourse through Re-Designing Button Labels

Authors: Stephen Samonte Tan (stephtan@vt.edu) Momen Bhuiyan (momen@vt.edu)

Advisors: Tanushree Mitra (tmitra@vt.edu)

Abstract: Effective deliberation is key to a democratic polity. Yet partisan intolerance in online discourse is on the rise. Powered by the disinhibition effect—a tendency to speak and act with less restraint online than one would offline—online conversations often devolve into streams of insults and epithets. For example, Twitter users with opposing political viewpoints were found to primarily direct insults at one another. Facebook users have reported the difficulty of co-existing online with people of different political opinions, often leading to reduced participation on the platform. Overwhelmed by incivility in comments, some news sites have even disabled commenting all-together, thereby restricting any form of online discussion. Unfortunately, recent research suggests that journalistic news values and reader political biases can compound the likelihood of incivility in online comment sections of newspapers. Since incivility has the power to alienate individuals from political discussion, it is important to explore ways civil deliberation can be supported, since democracy is premised on active citizenry capable of having civil discussions despite their differences. How can we enable people to co-exist online amidst differences in political opinions and interests? Can we design interfaces that help nudge people to engage in tolerant online discourse, despite their differences? The project proposes to experimentally test how citizens behave when they can click on one of the following five buttons, inspired by the “Moral Foundations Theory” (Graham, 2018), appearing next to articles and comments on news sites and social media platforms: "Like", "Respect", "Care", "Fair", "Incivil".

R10: Virginia Forest Visualization

Authors: Barrett Lattimer latb2014@vt.edu

Advisors: Chris North north@cs.vt.edu, Jobriath Kauffman jkauffma@vt.edu
(Conservation Management Institute, College of Natural Resources and Environment)

Abstract: Like all over the world, the forests of Virginia are constantly harvested for their wood. Concerns over a wood shortage in Virginia have been raised recently due to variables including rising European demand for Virginia wood, global warming, and plain competition alone between United States mills. A team of computer scientists and forestry majors from Virginia Tech constructed a simulation of Virginia forests over the next 30 years. Virginia's forests were simulated in a spatially precise manner keeping track of tree age, tree type, volume of wood by type, and growth rate for each type of wood. While having such detailed simulation results can be advantageous for identifying trends, the amount of data makes finding those trends over time difficult. As opposed to small mills and small landowners, large mills and massive landowners realize the many dangers facing Virginia's forests and have data readily available to help them make intelligent economic strategies for the future. Local mills and small landowners should not be put at such a disadvantage and should have access to the same cutting edge predictive forest models as the large mills and large scale landowners. I constructed an interactive online tool that presents the simulated data for Virginia forests in an intuitive and informative way. My tool consists of a drop down menu in which the user can select a variable they would like to observe by county, such as Average Age of Trees. When a variable is selected, a map of Virginia's counties is colored by its corresponding value for a given year. The year can be changed at any time by a slider. A plot is also generated, showing the change in each county for the corresponding variables throughout the 30 years simulated. I will be putting this tool up on a website and distributing the link to Virginia mills and landowners, allowing them to see this simulated data in an easy way and draw their own conclusions. This tool has already allowed for the simulation team to identify some forest trends throughout Virginia that wouldn't have been seen otherwise. In the future the Virginia Tech Forestry Department hopes to use this tool to visualize similar data for Georgia. This tool allows access to cutting edge simulation data with little to no technical knowledge, providing an interactive experience that keeps everyone informed on the predicted health of Virginia.

R11: OpenDSA eTextbook to Enhance Student Understanding of B+ Tree

Authors: Liling Yuan (yliling@vt.edu)

Advisors: Clifford A. Shaffer (shaffer@vt.edu)

Abstract: OpenDSA produces infrastructure and materials to support courses in a wide variety of Computer Science-related topics such as Data Structures and Algorithms (DSA), Formal Languages, Finite Automata, and Programming Languages. Its signature feature is collection of visualizations and exercises. Visualizations help students to understand each algorithm by watching the change it makes over a series of steps. Exercises let students consolidate what they have learned. The B+ Tree is one of the most complex but also the most important data structure for students to learn. This project is to provide better instructional methods for B+ Trees through a series of visualizations. To help students better understand the purpose of the B+ Tree, instructions show records with both keys and values. B+ Tree insert and delete operations are complex because there are a lot of different situations when we add or delete record, so the visualizations show each step, and explains the B+ Tree by how it changes using highlighting, affected records and showing their movements. Examples include add, find and delete in different situations. All the visualizations are written using HTML/CSS/JavaScript. They will be available to students beginning in Summer 2019.

R12: Algorithmic Composition in Virtual Reality_WEI

Authors: Weiqi Yuan wiki0831@vt.edu

Advisors: Sang Won Lee sangwonlee@vt.edu

Abstract: Music creation environment constructed with programming concepts, including I/O, logic operator, loop and abstraction will be built and used in the study. Users with different programming background will be invited to participate in this VR experience and a survey on their programming knowledge will be conducted before and after the experience.

C1: Consensus Clustering of Connectomes

Authors: Yevhen Pylypenko (yevhenp@vt.edu) Kaleb Jubar (ktjubar@vt.edu) Andrew Franco (ajf17@vt.edu) Kevin McCormack (kevjm19@vt.edu)

Advisors: T.M. Murali (murali@cs.vt.edu)

Abstract: Clustering algorithms are used to separate and classify graph nodes into clusters—groups based on a specified measurement of mutual proximity. That is, nodes considered to be “close” to each other are put into the same cluster, such that nodes in a different cluster are considered to be sufficiently “different”. Clustering algorithms often differ in input requirements, how they define and distribute nodes into clusters, and what kind of output they produce (as well as the analysis tools available for such output). This project aims to apply multiple clustering algorithms on brain connectome graphs in order to reach a consensus—i.e., analyze how the clusters produced by each algorithm compare and differ from one another, and determine the pros and cons of running each algorithm. Thus, we hope to discover whether one particular clustering algorithm performs best for a particular connectome due to its inherent graph properties, and future research can expand on this idea.

C2: Multi-Feature Analysis of Research Papers to Classify Retracted Studies

Authors: Jonathan Alexander, jma@vt.edu Raghu Srinivasan, raghu@vt.edu Henry Wang, henrymw@vt.edu

Advisors: Tanushree Mitra, tmitra@vt.edu

Abstract: Scientific studies are a mainstay of current mass media. The problem we aim to tackle is misinformation and misconduct in these studies. Many of the studies cited by popular media are fraudulent and can mislead the public. We define misinformation as data or conclusions shown in published studies that was created through acts of scientific misconduct. Some examples of misconduct include data manipulation, invalid conclusions, cherry picking samples, and flawed scientific design. The common recourse for studies that take part in these methods involves retraction by the journals review board. However, this does not stave media outlets from making claims based on them, or solve the headlines that have already been written. This can lead people to make poor decisions based on these invalid articles and the headlines they see because of them. The difficulty in identifying misleading scientific studies is one of the main reasons for the wide spread of misinformation. It is difficult to determine if a researcher manipulated or cherry picked their data before submitting the paper to the journal. There is also a subset of journals labeled predatory journals that lack a true review process in order to garner higher publishing costs from researchers. Such journals take payment from researchers or other groups that simply want their study put out there even if it is not of scientific merit; these predatory journals may also not retract studies that are invalid making it even harder to detect fraudulent studies. Furthermore, in depth reviews of the methods and data of a study are done by humans and is a very lengthy process, introducing human error and leading to slow response time retracting studies that slip through the process. Our methods involved measuring several features of a scientific study from a sample of both retracted and nonretracted papers, analyzing these metrics, and then feeding these features into a learning classifier. After comparing results from training several classifiers, we obtained a final accuracy of ~60% compared to a baseline of 50%. We found that linguistic features (such as words per sentence, number of six-letter words) ended up having a more significant impact on the classification, whereas sentiment features (such as positive, negative, and neutral sentiment scores) did not have much of an impact. We can therefore conclude that it is possible to to classify retracted from nonretracted articles with a sufficient accuracy from analyzing numerous types of features.

C3: Tourism Destination Websites

Authors: Matt Crawford (mattc97@vt.edu) Sotirios Nicholakos (snichola@vt.edu) Robert Rizzo (robrizzo@vt.edu) Viet Doan (viet96@vt.edu) Jackson Salopek (jsalopek@vt.edu)

Advisors: Ed Fox (fox@vt.edu) Florian Zach (florian@vt.edu)

Abstract: In the United States, cities like Philadelphia, Pennsylvania or Portland, Oregon are marketed by destination marketing organizations (DMOs) to potential visitors. DMO websites mostly are a collection of information on attractions, accommodation, transportation and events. Dr. Florian Zach seeks to conduct a historical analysis of the outbound and internal links of DMO websites in order to learn which website DMOs point to and how this changes over time. We will assist him in this analysis by analyzing data from top tourism websites from each state in the United States, 98 in total. Dr. Zach's expected impact of this research is the ability for tourism destination decision makers to learn past trends among DMOs' outbound links and potentially determine the future direction of destination online marketing. From the DMO data, we will deliver a set of matrices from which Dr. Zach can perform his own analysis in Python. Dr. Zach expects DMO websites to point towards a set of common websites like social media sites, Tripadvisor, Yelp, local websites, etc

C4: An Analysis of the Sexist and Racist content in Gaming Channels by Genre

Authors: Kyle Czech ckyle96@vt.edu Dat Bui deweybat@vt.edu Tucker Crull tucker24@vt.edu

Advisors: Tanushree Mitra tmitra@vt.edu

Abstract: Our group set out to analyze the problem of hateful and offensive content on online gaming streaming services. The problem is that people hiding behind screens can be very hateful to the other people. This is especially true in the world of online gaming for our personal experience. To look and analyze this problem our group chose to look at Twitch and YouTube as they are the most common and most used gaming streaming services in the USA. Our goals for our project is to look at how hateful and offensive content are used across gaming genres, develop a way to predict hateful and offensive content in the chat logs of different streamers, look at how the different streaming services compare to each other across the different genres of gaming, and to try and prove that the sex of the stream directly affects the amount of hateful and offensive content they receive. Our motivation for this study comes from our love of gaming and consuming gaming content. Combine that with the fact that gaming and e-sports industries are some of fastest growing industries in the world. This study could be extremely useful in finding ways to stop hateful people who use hate and offensive content and ruin the streaming content for other views. Also, our study would be very informative about how sex bias come into play when people uses hate and offensive content online.

C5: Louisiana French

Authors: Joe Chu (jchu247@vt.edu) Will Duong (wduong@vt.edu) Abhi Oddula (oabhi@vt.edu) Rahul Ramki (rahul423@vt.edu) Saharsh Shrivastava (saharsh2@vt.edu)

Advisors: Katie Carmichael (katcarm@vt.edu) Aarnes Gudmestad (agudmest@vt.edu) Edward Fox (fox@vt.edu)

Abstract: The goal of our project is to assist in the sociolinguistic research of Dr. Carmichael and Dr. Gudmestad. Their project involves the research of Louisiana French, a dialect of the French language, spoken in Louisiana. When we first contacted our clients expressing our interest in this project, they sent us a corpus of interviews conducted by Dr. Carmichael. These transcribed dialogues analyze the dialogue of the interviewees speaking in Louisiana French. Our group's goal is to help our clients analyze verb frequencies in the interviews given to us. We are initially given a verb bank, and our responsibilities are to find and count each conjugation of each verb in the bank. To find each conjugation, we had to account for both regular and irregular verbs in the bank. Irregular verbs are conjugated differently, and for that reason, it involves a different conjugation process. Our first phase of the project involved getting familiarized with the French language in general and how conjugations of a verb, in French, take place. During this time, we have also decided to use Python to build scripts that would extract data from our corpus. Though generating Python scripts was original implementation, we have soon realized that using TreeTagger software would be the best and most efficient way to go about the next steps of our project. TreeTagger is a software that can annotate text with part of speech information. Using this software, we are able to calculate the frequency of specific verbs and verb conjugations by parsing the data returned from TreeTagger. TreeTagger uses parameter files for specific languages in order to determine the parts of speech for sentences. Our group is using the French parameter file. We have progressed towards a final deliverable for our client as well as a user manual for future use. If they do decide to add data to their interview corpus, they will have an understanding as to how to run their data with TreeTagger.

C6: Street Art: an immersive VR experience

Authors: Carlisle Hughes carlih5@vt.edu JooYoung Whang joo918@vt.edu Yasmine Belghith byasmine@vt.edu

Advisors: Steve Harrison srh@vt.edu

Abstract: Description/Background: Street Art: an immersive VR experience, is a virtual reality application where users are able to create worlds in the form of block-built environments before spray painting them to their liking. Methodology: In order to implement this application successfully, we first needed to produce a proof of concept. The latter would demonstrate the feasibility of our project and verify that the multitude of features we intended to construct and shape together would create a harmonious and intuitive immersive gaming environment. We adopted the Oculus Go as our VR platform of choice because of the ease of development, its availability, as well as the different assets and open-source libraries we could include in our project. Different components were achieved separately, using Unity and the Oculus Go Unity Integration package, and tested on the Oculus Go. After testing, each functionality was reviewed and optimized to allow for a more efficient final product. Building (adding and removing blocks), painting with vertex colors (spray paint effect), mobility (teleportation within the VR environment), as well as the user interface (gaming actions menu) were merged together as a prototype on which evaluation was performed. Keeping various scenarios in mind, we put ourselves in the user's position and tried all the game functionalities. We are currently working on the refinement of Street Art: an immersive VR experience before presenting it to the public. Uses and Platforms: We believe this application has substantial creative potential. Springing from the theme "Art as a way to organize the world", it brings together architecture, a staple in many cultures and in the arts, as well as painting, another artistic and creative staple. This could also be the foundation for more development, by constantly improving it, refining it and adding new functionalities such as multiplayer. Future work could also include spreading this application to other VR platforms as well. Our user base is wide as this application has neither an age, nor an ability limit. This application could be a testament to the artistic essence of the community.

C7: Expanding HEART using Advanced Advertising Techniques

Authors: Juan Segura (juan97@vt.edu) Phillip Hrinko (philh97@vt.edu) Keller Han (kelhan18@vt.edu) Vanessa Lomeli (vanessal@vt.edu)

Advisors: Scott McCrickard (mccricks@cs.vt.edu) Abby Steketee (abbymys@vt.edu) Samantha Harden (sammadee@vt.edu)

Abstract: Healthy Eating and Active Readiness for Two (HEART) is an educational intervention program that helps pregnant women prevent excess gestational weight gain. The program is a research study under Dr. Samantha Harden that meets with a group of women, giving them the resources to maintain a healthy lifestyle throughout their pregnancy. Throughout the initial meetings with Dr. Harding's PhD student, Abby Steketee, it was clear that the HEART study was beginning to have issues recruiting women for the study since they were moving away from recruiting through the Carilion Clinic. For our Human-Computer Interaction capstone project, we decided that we could best support this study by creating a recruitment website. Through our initial meetings with Mrs. Steketee and Dr. McCrickard, we were able to determine the essential information to be included in the website and what the best way to present that information would be for pregnant women. Using a Heuristic evaluation, we were able to receive some initial feedback for our website and we were able to correct a few heuristic violations in our initial website draft.

C8: Picturing the Trails

Authors: Nicholaus Clark ntc0531@vt.edu Dylan Finch dkfinch@vt.edu John Kook johnk97@vt.edu Emily Maher emilym97@vt.edu

Advisors: Michael Horning (mhorning@vt.edu) Scott McCrickard (mccricks@cs.vt.edu)

Abstract: Picturing the Trails is a web application that helps users to create a photo story with a given picture sets. Currently, people who are interested in creating a story must go through photo collages to pick out the perfect photos to use. Many people struggle to choose the perfect photos to make the perfect stories. Others simply do not have the time to create the story after a hike or a family activity due to life responsibilities. With Picturing the Trails, anyone can create a professional photo story by simply uploading the photos to the application. This way, you can organize the photos post-hike even during your busy life schedules. This application will store the other stories that you have created as well, naturally making a magnificent, organized photo collage.

C9: BarkLight: Ubiquitous Computing in a Naturalistic Environment

Authors: Connor Smedley csmedley@vt.edu Ryan Jakiel rjakiel@vt.edu Puriwat Lahpong pulahpon@vt.edu

Advisors: Matt Wagner wagner.matthewdavid@gmail.com Scott McCrickard, mccricks@cs.vt.edu Tianyu Ge tianyu@vt.edu Shuo Niu shuoniu@vt.edu

Abstract: Advances in technology continue to give us amazing new devices that assist us in our daily lives. However, these devices are demanding more and more of our attention, and we can spend most of the day looking down at our wall displays and smartphones and smartwatches, blinded to the world around us. Ambient displays have the potential to convey information to users without commanding undue attention from them. They naturally fit into their environment and display information that is intuitive to what the user is looking at. BarkLight is a new type of ambient display we are constructing with a flattened piece of tree bark that we fitted with over 1100 LED's. We are using the LED's to display animations that show the current weather conditions in a peaceful, uninterruptible way. BarkLight centers around principle ideas of ubiquitous display heuristics. Although BarkLight is an interest piece that will initially draw a lot of attention to itself, it will be a great ambient weather display as it blends into the users' usual environment.

C10: UI/UX Improvements to MatWeb

Authors: Anthony Nauls (nauls97@vt.edu) Nayan Gupta (nayan.gupta@vt.edu) Alex Medovar (alexm97@vt.edu) Phillip Hrinko (philh97@vt.edu) Abhi Oddula (oabhi@vt.edu)

Advisors: Rich Charles (rich08@vt.edu)

Abstract: MatWeb is a website that contains a searchable database of material properties including data sheets for various types of engineering materials. This database is currently over 130,000 data sheets in size including metals, plastics, ceramics, composites, etc. While this website has been around for many years and is reliable in what it does, its design and UI have not been updated in about ten years. For our Software Engineering Capstone, we decided to take on this project with the goal of improving some of the UI and design aspects of the website. Since the current website has not been updated in years and most users have been accustomed to the current look of the website, a challenge presents itself to update the look of many pages while also not moving around functionality. This means that changes done to the website should not be large in scope so current users do not become lost, but are sufficient enough to give the website a more modern look for newer users. With this in mind, our project group needed to decide what aspects of the website needed the most improvement, while also ensuring that the overall feel of the website stays relatively similar. After planning and creating mock-ups, we identified the changes that we planned to make during our development period. These would include changing the navigation bar, footer, homepage, and other various pages on the website that were cluttered with text or simply outdated. While many of these UI improvements may appear simple, the challenge of working with an old code repository presented itself and made us adapt to the situation. Furthermore, due to the large scale of the database and power of its many search options, it was important to not alter any of the functionality to the website. By adding our UI and design improvement to the website, we hope to allow newer users to return to the website more often along with retaining the number of current users on the website.

C11: Healthy Hokies

Authors: Griffin Norris - gn196635@vt.edu Parker Robinson - parkr96@vt.edu Sarah Marcus - stmarcus@vt.edu Tracy Pan - tracyp6@vt.edu Nick Onofrio - nicko5@vt.edu

Advisors: Rich Charles - rich08@vt.edu

Abstract: A common frustration of students that use the facilities at War Memorial Gym and McComas Hall are the crowds that are often found there. Many students end up spending more time at the gym waiting to use a facility, than they do actually using the facility themselves. One of the main causes of this problem is that there are many students that use these facilities but students are not aware exactly how crowded the gym is before they go. HealthyHokie is a website designed to increase the ease of usage of on-campus gyms, War Memorial Gym and McComas Hall, at Virginia Tech. When a student goes to one of these gyms, they can use HealthyHokie to check into the facility from a webpage. When the student leaves the gym, they can check out. If the student forgets to check out, HealthyHokie will automatically check the student out after 90 minutes. As students check in and out of the gyms on HealthyHokie, the total number of students currently at the gym is recorded. This total number of students is displayed on the HealthyHokie web page allowing students to use the web page to see how many other students are at the facility, before they even check-in to the facility themselves. The functionality of HealthyHokie aims to both increase facility usage and increase facility efficiency. By helping students choose times to attend the facility that are less busy, the facility will be less crowded making it more enjoyable to be at and it's facilities easier to use. A future extension of HealthyHokie will reward students for checking in to the gym on a point based system. This feature will encourage students to attend the gym more frequently by providing incentives. Virginia Tech repeatedly makes its way to the top of the rankings for things such as "The Fittest Colleges in the U.S." and "Healthiest Colleges". This application has the potential to keep Virginia Tech high in these rankings by helping Healthy Hokies become healthier.

C12: Tree Viz

Authors: Conner Caprio caprioc@vt.edu Austin Zensen auze217@vt.edu Moira Pelton moira15@vt.edu

Advisors: Ed Fox fox@vt.edu

Abstract: The TreeViz project was developed as part of the capstone course CS 4624 Multimedia, Hypertext, and Information Access and was completed over the course of the Spring 2019 semester. The goal of this project is to allow students taking the CS 2114 data structures classes to visualize code written for tree data structures through CodeWorkout homework exercises. Students will be able to see the changes their code makes to a tree data structure in real time as they click through visualization steps in the existing interface and our changes to accommodate tree data structures. This project is an extension of the client, Mostafa Mohamed's existing work to create visualizations for linked list data structures for exercises for the CS 2114 and 3114 data structures classes. This project was implemented in two main phases, with the first goal of generating a tree visualization from student code and the second goal of being able to visualize changes to the tree data structure as the student code executes. We wrote a set of JavaScript files for the OpenPOP (OPEN Pointer OPERations) server that is used by OpenDSA and CodeWorkout to generate and parse the execution trace of student code for changes to the underlying data structure. Our solution analyzes the execution trace of the student's tree code and determines the pointers created during execution to build a replica of the tree data structure being created in memory. With each step in the execution, our code will determine the pointers that make up the tree and draw the tree using the JSAV library to provide the physical representation of the nodes and branches within the tree data structure. This way, the student can view and step through the execution of their code and visualize the changes to their tree data structure.

C13: BirdFeed

Authors: Chandler Manns - chandm8@vt.edu Stephen Tewes - stewes36@vt.edu

Advisors: Timothy Stelter - tstelter@vt.edu Scott McCrickard - mccricks@vt.edu

Abstract: Bird watching, or birding, is an outdoor activity where people engage in observing birds. We present BirdFeed, an educational Android mobile application that is geared towards a younger generation of bird watchers to build a personal collection of pictures, audio, and notes of the 25 most common birds in Virginia. BirdFeed is designed to be an easy-to-use data collection, field guide, and reflection tool for new birders. We worked with the Seek Education, Explore, DiScover (SEEDS) program which aims to educate children while supporting their exploration and discover different aspects of nature. On this poster, each phase of the human-computer interaction lifecycle for BirdFeed is concisely presented.

C14: Adult Day Services Memory Masterclass Promotional Video

Authors: Maddie Kulik, maddie14@vt.edu Pablo Castillo, pcast11@vt.edu Jose Zurita, jose95@vt.edu

Advisors: Edward Fox, fox@vt.edu Ila Schepisi, schepisi@vt.edu Sara Amos, sara17@vt.edu

Abstract: The goal of this project was to create a promotional video for Virginia Tech's Adult Day Services center, specifically to advertise for their Memory Masterclass program. Adult Day Services is a center located within the Human Development and Family Sciences Department at Virginia Tech. They are licensed by the Department of Social Services to offer personal care, health monitoring, meals, therapeutic activities, dementia care, and recovery assistance. They serve typically 18 participants each operating day who average about 75 years of age. According to ADS's mission statement, the center is dedicated to providing a center focused on the well-being and optimal functioning of its participants, a resource for caregiver support, an education opportunity for students, and a community among generations of children, college students, and adults. One of ADS's main service offerings is their Memory Masterclass course. This course is offered in 6-week sessions to participants over 55 years of age who want to maximize their brain health. The focus of the course is to educate and serve people who have been diagnosed with Mild Cognitive Impairment (MCI). MCI is not a symptom or precursor to Alzheimer's or dementia, but rather a condition that occurs as aging changes brain function. In the 6-week course participants learn strategies for application to daily life that can strengthen brain reserve as you age and get connected with others who have similar concerns about memory. Our main objective was to create a promotional video that Adult Day Services could use on their website to inform and attract people to take the class.

C15: CSRC Interview Check-in System

Authors: Teddy Todorov, ttodorov@vt.edu Nick Gill, nsgill1@vt.edu Ben Austin, bennoxys@vt.edu Joe Chu, jchu247@vt.edu Travis Rieger, trieg@vt.edu

Advisors: Rich Charles, rich08@vt.edu

Abstract: In previous years, the VT computer science department has been using pencil and paper to keep track of Career Fair interviews. This means that employers would have to walk up to that sheet of paper and look through it to see if their interview candidate had checked in. Furthermore, after the event, a CS@VT employee would need to manually consolidate information such as number of students interviewed from that piece of paper. Our goal was to create an online interview check-in system to streamline the process for students, employers, and faculty related to the CSRC. We aimed to keep the interview check in process simple for students, while still allowing Virginia Tech employees to view statistics about individual career fairs and companies to see a list of all of their checked in candidates. Our website was created with three different simple systems designed for the people we envisioned using it. The default view when you visit the website is for students, and is a four screen loop that prompts the student to swipe their student ID and select a single company they are interviewing with, then offers a review screen and a thank you before automatically going back to the start of the loop. The admin page for VT employees is a single page with information about current and past career fairs, including the companies that attended and statistics about the number of check-ins, the number of unique students that checked in for an interview, and the companies that gave interviews to VT students. The third view, the employers page, is intended to give companies a schedule for interviews by showing them a list of each student that has checked in for an interview with them and displaying the time of their check-in next to the students name.

C16: Fostering Digital Cooperative Management on the Appalachian Trail

Authors: Daniel Ocheltree djo96@vt.edu Parker Irving parker45@vt.edu Campbell Johnson campbel1@vt.edu

Advisors: Christina McIntyre cmcintyre@vt.edu Scott McCrickard mccricks@cs.vt.edu

Abstract: The Outdoor Club at Virginia Tech is responsible for maintaining over 40 miles of the Appalachian Trail. The Club faces many challenges with maintaining this section of trail including lack of public reporting, scheduling issues, work trip reporting, and time management. We sought to build an application that would address these issues. Our web application allows for the public reporting of trail issues, which in turn are available for OCVT leaders to review and address. OCVT members have the ability to schedule work trips and create new tasks, all while maintaining a record of past, present, and future work. Currently, if you are a hiker and discover a trail issue on the OCVT maintained section of trail, there is almost no way for you to report that issue. The club relies on many volunteers who hike the trail often, as well as a shared trail knowledge amongst club leaders. However, one of the big issues the club faces is member turnover, since there is new leadership every year it is hard to maintain that shared knowledge. We learned of these issues by interviewing the faculty advisor for the OCVT, Christina McIntyre. We also consulted the practices of other trail clubs, and the Appalachian Trail Conservancy to identify any existing solutions, but in our research, we couldn't find any prior work that addresses our identified problems. After analyzing the problem space we decided to move forward with a two-fold solution, a public-facing system that allows hikers to report issues they encounter on the trail, as well as a trail management system which organizes these issues and helps the volunteers create, plan, and execute work trips. We started our design by creating wireframes and getting feedback from hikers, our client, and fellow classmates. Then we created a skeleton website and got more feedback from the same individuals. We wanted the site to be very mobile friendly and the feedback from our peers helped us create a better mobile interface. We continued to iterate on the design and add new features. During this phase, we were in constant communication with our client Christina and got some great usability feedback. For our analysis phase, we handed over the application to our client and the OCVT. We also distributed signs to put on the trailhead which explains how to use our application. The OCVT is using our app to help them with their upcoming work trips and we are getting feedback from real-world users. Right now is peak thru-hiker season on the Appalachian Trail and the OCVT's section of trail has about 20 hikers per day coming thru, and those hikers now have the ability to use our application to report trail issues they encounter.

C17: Department of Computer Science Absence Survey System

Authors: Harrison Gertler - hsg32@vt.edu Sami Morency - samim7@vt.edu Ernesto Cortes Groman - ernestoc@vt.edu John Clarke - cjohn1@vt.edu Kevin Shah - krs2070@vt.edu

Advisors: Rich Charles - rich08@vt.edu Tabitha Hartman - hartt@vt.edu

Abstract: The goal of this project was to redesign the system that the Department of Computer Science at Virginia Tech uses for recording and disseminating personnel absences. Our project allows users to Create, Read, Update, and Delete (CRUD) personal entries through a web application. Our web application replaces the current system of a Google Form that the Department of Computer Science currently uses to record absences. Our team created a web application where users can view their entered absences on the home dashboard. From this dashboard they can edit or delete these entries. To create a new entry, they use the Absence Survey which auto-fills their Name and Employee Type. They select a Type of Leave from a dropdown menu which populates the rest of the survey based on the choice they select. Their entries are stored in a MySQL database which populates the dashboard. Our project gives users two methods of viewing upcoming weekly absences: an automated email sent out on Monday mornings, and a weekly dashboard that is viewable through our web application. If a user edits or deletes their entry, an email is sent out the next morning with any and all updates. Our team also created an Admin Panel for specific users to be able to view, edit, and delete all entries, sort and filter the entries, and export the table as a CSV file. We seek to expand the capabilities of our project to give users and admins more features such as filtering of on the home dashboard and weekly dashboard.

C18: Functional Subgraphs in Brain Networks

Authors: Dallas Viar - dallasv6@vt.edu Ethan Hoepfner - ethanmh@vt.edu Luke Wolff - Lwolff@vt.edu Ryan Frodge - frodger@vt.edu

Advisors: T. M. Murali - murali@cs.vt.edu

Abstract: Brain networks have the potential to provide insight into the interactions and functionality of human behavior and control. With this capstone project we aim to identify correlations in brain interactions and use the information to divide brain networks into discrete subgraphs using quantitative techniques. These subgraphs are useful in analyzing interactions between different physical areas of the brain, and may help lead to a better understanding of these complex interactions. We use non-negative matrix factorization (NMF) on fMRI BOLD data from twenty eight subjects. By using a non-negative matrix factorization, we obtain a parts-based decomposition of the entire network. This allows us to more easily inspect the resulting matrices for correlations, and identify and differentiate subgraphs. This was important due to the noise as well as the sheer volume of data we were attempting to analyze. Our methodology identifies a number of distinct networks within the brain, separated according to the temporal patterns of expression in the brains of the subjects during the tests. The fMRI data was gathered while subjects were performing two types of cognitive tasks: a Stroop interference task and a local-global perception switching task using Navon figures. Each task is performed with high and low degrees of cognitive demand. We then investigated the ways these subgraphs are activated during the various tasks, at each level of cognitive demand. This project attempts to replicate the published results of a project by Danielle Bassett's group. Their results showed that the degree and order of activation of subgraphs changes alongside cognitive demand. This allows information to be determined about the changes from one cognitively demanding brain state to another.

C19: Sustainability Values: Interactive Website for Values Diagnostic Reporting and Analysis

Authors: Mary Haninger, maryh15@vt.edu Leah Triplett, jtleah7@vt.edu Will Miksovic, wmik9@vt.edu Andy Paulonis, paulonis@vt.edu Kaleb Jubar, ktjubar@vt.edu

Advisors: Edward Fox, fox@vt.edu Bruce Hull, hullrb@vt.edu

Abstract: The goal of this project is to develop a website that supports a survey created by our client, Dr. Hull. This survey is designed to help sustainability professionals and students learn about their own preferences, values, and opinions as well as how they compare to others around them. This will in turn help them understand themselves, become better listeners, build productive collaborative efforts, and manage conflict. This website must allow users to take Dr. Hull's survey, analyze their answers, and provide them with feedback about where they fit in relation to their sustainability values. After giving the user their results, all the data needs to be compiled in such a way that Dr. Hull can download and analyze it for his own research. All of the data accumulated will be kept anonymous to maintain the user's privacy. This data will allow Dr. Hull to see where his students' and colleagues' values lie and how to tackle these sustainability issues in the future. This project was started by a CS 4624 group in 2018 (Brizuela, X., Stewart, C., Mistry, H., & Eltepu, S.). This group's effort is being used as the base for our current project.

C20: Generative Mathematical Models for Real World Connectomes

Authors: Lucas McCormick mluke16@vt.edu Samuel Furman sfurman@vt.edu Conor Kelly conork@vt.edu

Advisors: T. M. Murali murali@cs.vt.edu

Abstract: A variety of generative algorithms seek to model brain connectomes, with the goals of recreating mechanisms that produce experimentally derived graphs. However, each model has its own strengths and weaknesses relating to how well it can generate realistic brain graphs. In this project, we studied and implemented several examples of two classes of generative models. Embedded models, which are constrained by geometric coordinates, and non-embedded models, which are not. We parameterized these models using empirical brain graphs in different organisms with real neural data found via neurodata.io and mappings determined from the Human Connectome Project's brain scans. The resulting networks were then analyzed for the following diagnostic factors: diameter, average shortest path length, clustering coefficient, degree mean, degree standard deviation, and degree median absolute deviation. In attempting to identify generative models that best capture the properties of the empirical graphs, it became apparent that no single model used in this study is capable of approximating real data across the entire suite of diagnostic statistics. It is often the case that most closely matching model for attributes like average shortest path length and clustering coefficient fails to come close for attributes like degree median absolute deviation. This result corroborates findings in the literature that ascribe the phenomenon to the many neurobiological mechanisms that drive brain development. Furthermore, applying the DDG probability function from the previous work to the human connectome saw a fit of $r^2=.96$ but failed to capture some key nuances found in the connectome's probability curve. Several other functions including the damped sinusoid and second order response function were fit against the empirical graph with less success. In the future more models must be investigated to see if there are any which can capture the diagnostics of the empirical data, without any tradeoffs. Empirically parameterized models such as the DDG model may provide a good starting point, but more work must be done to determine underlying functions which match real world data.

C21: Accessibility Map @ VT

Authors: Mahira Sheikh (smahira8@vt.edu) Yasmine Belghith (byasmine@vt.edu) Jack Danisewicz (jackd1@vt.edu) Josh Wenger (joshnw@vt.edu)

Advisors: Hani Awni (awni1@vt.edu) Scott McCrickard (mccricks@cs.vt.edu)

Abstract: Description/Background Our project, Accessibility Map, aims to provide smart walking routes around the Virginia Tech Campus that take into account accessibility barriers.

Methodology To understand our problem more clearly the first thing we did was gather information from officials that work with disabled students as well as accessibility standards around campus. After interviewing a number of Virginia Tech officials we recorded work activity notes and organized those notes into a work activity affinity diagram. The purpose of this was to more clearly understand the needs and wants of campus officials regarding campus accessibility. From here we developed personas, scenarios, and design requirements. Using these we came up with a number of low-fidelity prototypes. We are now in the process of developing a high-fidelity prototype that includes features such as custom routing based on a user's entered abilities and barrier reporting along a route.

Uses and Platforms We believe this application will have many uses. Our main user case will allow disabled users to enter their accessibility constraints, such as wheelchairs, walkers, and audio sensitivity, and navigate from point A to point B on campus with a route tailored to them. In addition, advocacy groups such as the ADA office will be able to use the application to view all of the permanent and temporary barriers in order to find out what barriers need fixing and allocate necessary resources to them. Accessibility Map will be available on all platforms. The web application is designed to be mobile as well as desktop friendly. Our application is being designed in order to empower a user in whatever situation they might find themselves in.

C22: PrepLab

Authors: Andrew Bond dbond07@vt.edu Steven McKim crowned1@gmail.com Nathan Kendrick knathan2@vt.edu

Advisors: Edward Fox fox@vt.edu

Abstract: PrepLab is a proposed online repository to aid instructors in the remixing and reuse of instructional materials (e.g. syllabi, assignments, learning activities, facilitated discussions, etc.). To facilitate easy adoption it was designed to be as simple as possible for content creators to upload existing material. This project set the constraints of what content may and may not be easily included. When uploading content, users shall select what creative commons license they wish to apply (the system may impose a minimum license) and confirm that they are the original content creator and/or own the rights to any material they are submitting. Instructions to content creators for structuring new material is also included. Upon initial upload of an item the results of the parser/indexer shall be displayed to the user who is given the opportunity to confirm/edit the results. This project is dedicated to working specifically on the Graphical User Interface (GUI) for the PrepLab. React.js, a front end javascript framework, was chosen as the design method upon which the front end of PrepLab is built. Other design elements added to the website include Bootstrap as well as traditional CSS libraries. Amazon Web Service (AWS) is used to host the website until a more suitable replacement can be determined when the other sections of this project are completed. As the backend pertains to a different part of the project, the database holding the instructional materials that will populate the site will be covered in future projects related to PrepLab. The other section that is not covered in this part of PrepLab will include parsing and indexing of documents as well as including an API in order to increase functionality and search capabilities.

C23: Cranial Genetics: Investigating Clustering Algorithms to Decode Gene Expression Data from the Human Brain

Authors: Arianna Krinos (akrinos@vt.edu) Luke Quinn (lquinn@vt.edu) Marcus Milks (msquared@vt.edu) Joshua Detwiler (joshde@vt.edu)

Advisors: T.M. Murali (murali@cs.vt.edu)

Abstract: In this capstone project, we clustered gene expression data to reveal properties of single human brain cells. Our goal was to investigate whether we could learn more about the interaction and function of elements of the human brain. To this end, we had a dataset from the Allen Brain Institute, consisting of gene expression data from a series of samples from the middle temporal gyrus of the human brain. Because the dataset contained a large number of cells and genes, and a number of genes that were often not measured, we needed to reduce the dimensions of the dataset to be able to work with it easily. We used Principal Components Analysis (PCA) for this task. Next, we analyzed the results using two independent approaches: (1) t-distributed Stochastic Neighbor Embedding (t-SNE), an algorithm used to project high-dimensional data to few dimensions in order to visualize relationships among those data, and (2) PhenoGraph, a method developed to cluster single cell expression data (Levine et al., 2015). We combined the results of these approaches into a 2D plot, where each cell corresponded to a point whose coordinates arose from the two most significant vector components, as decided by the t-SNE algorithm. The points were colored by their clusters as determined by PhenoGraph. We saw generally good correspondence between the topographical location of the cluster via t-SNE and the cluster color via PhenoGraph. Finally, we performed hierarchical clustering (hclust in R) to demonstrate clustering of samples based on the cell type within which they were measured by the Allen Institute. We found that we were able to get two well-defined clusters of neuronal vs. non-neuronal cells. We also envision potential future applications of correlating our gene network analysis to protein interaction network data to more accurately describe brain function.

C24: Navigation Routing in Brain Networks

Authors: Sean McHugh seanfmch@vt.edu Barrett Lattimer latb2014@vt.edu Kendra Song ksong15@vt.edu George Gavrila agavrila@vt.edu

Advisors: T.M. Murali murali@cs.vt.edu

Abstract: Navigation Routing in Brain Networks Kendra Song, Barrett Lattimer, George Gavrila, Sean McHugh Advisor: Dr. T. M. Murali Our goal was to evaluate the performance of a plausible method of communication in brain networks called Navigation Routing. Communication within any network requires frequent path computations between source and destination nodes. However, many shortest-path algorithms, such as Dijkstra's, are centralized, meaning that routes are determined by an external source that possesses information about the entire graph. It is more plausible that communication in the brain happens in a decentralized manner, where routes are computed using only local knowledge stored in individual nodes. We investigated one decentralized path-computing algorithm known as Navigation Routing, proposed by Seguin et al. (PNAS, 2018). We implemented the algorithm and replicated the findings of this paper. Our implementation ran on the same data set used in the paper: a publicly available connectome (Glasser et al., Nature, 2016), or mapping of the connections in the human brain. We compared the paths produced by navigation routing to actual shortest paths found using Dijkstra's algorithm, and saw that navigation routing produced near-optimal paths. Additionally, we found that both randomizing the position of nodes and randomizing the edges of the graph greatly reduced the effectiveness of navigation routing. These findings corroborate those by Seguin et al. and suggest that there are spatial and topological properties of brain networks that facilitate navigation routing in finding efficient communication paths in a decentralized manner. It was not always clear what methods were used by the authors to arrive at figures in the paper, so we frequently made judgement calls in what techniques to employ. This challenge deepened our understanding of Navigation Routing and computational methods used to study networks.

C25: The Impact of Reddit's Coverage of the 2018 Parkland School Shooting on the Political Polarization of Online Communities

Authors: Sourav Panth - Spanth@vt.edu Chris Blair - chrisb56@vt.edu Numan Khan - numank@vt.edu Bright Zheng - bzqzheng@vt.edu

Advisors: Tanushree Mitra - tmitra@vt.edu

Abstract: Our research project on the impact of reddit's coverage of the 2018 parkland school shooting focused around one of the most deadly high school shootings in US History, the Stoneman Douglas School Shooting in Parkland, Florida on 2/14/18. School shootings have been a focal point of politics and a large risk to students recently. The research questions that we set out to answer include what were the prevailing topics discussed after the shooting, how do different groups perceive the issue of mental health in response to Parkland, over time how did the political affiliation and sentiment change, and finally how did sentiment and political affiliation affect the swing of online discussions?

C26: Online Role-related User Classification on Twitter

Authors: Gregory Pickett spgreg7@vt.edu Kenneth Worden kworden@vt.edu Adam Wilborn adamw97@vt.edu

Advisors: Liuqing Li liuqing@vt.edu Edward Fox fox@vt.edu

Abstract: The main goal of this project was to provide a web application that will host the existing TWIROLE model. TWIROLE, a hybrid model for role-related user classification on Twitter, which detects male-related, female-related, and brand-related (i.e., organization or institution) users. TWIROLE leverages features from tweet contents, user profiles, and profile images, and then applies the hybrid model to identify a user's role. The main use of TWIROLE is to aid future evaluation efforts and research studies relative to investigations that rely upon self-labeled datasets. The web application is made to be easy to use and navigate allowing for a wide range of audiences including researchers or common twitter users. Other goals of the project were to add a new classifier to the model that will improve the accuracy of TWIROLE. The model originally had an advanced feature that used the k-top words method to analyze users and classify them. Essentially, looking at all of a user's tweets and ranking the k top words used and classifying a user based on what the word is. The final product of the model also uses emojis to classify a user but instead of k-top words we check the frequency that an emoji is used. In addition to this feature another feature was added that tracks the second or third person score. Once features were added, the model was then trained again on the existing data set improving the accuracy. The website is made up of an HTML page using react on the front end, the backend (TWIROLE) is made using Django to render the html page, host images and other resources and expose a GraphQL API. The front end makes AJAX calls to the GraphQL API which obtains the information that will be displayed on the website. The website is aimed at being simple to manage and update with experience in web development specifically Django, React and GraphQL. The website is hosted by DLIB or Digital Library Research Laboratory using their domain name.

C27: The Effect of News Corrections in Today's Online Information Sharing

Authors: Nathan Gibson nathanag@vt.edu Kibur Girum kibur1@vt.edu Jacob Benjamin jbbenj@vt.edu

Advisors: Tanushree Mitra tmitra@vt.edu

Abstract: Fake news and the spread of misinformation are increasingly prevalent sources of concern in the age of the internet and information technology. We aimed to study the connection between news corrections and the spread of misinformation. We collected correction-related tweets and compared them across different news organizations. We also collected correction data and correction article comments from the New York Times to find unique characteristics related to corrected articles, as well as the sentiment of readers regarding corrected articles. Our results suggest that liberal news outlets share more correction tweets than moderate or conservative outlets, and corrections are usually retweeted at higher rates than normal tweets. Many corrections are made by the New York Times, but very few are shared on Twitter. We found that the topics of corrected articles differ from the topics of articles published by the New York Times overall. In addition, we did not find evidence that readers react less positively to corrected articles. We hope that our work sheds light on the effect of corrections in the spread of news.

C28: PolitiGauge

Authors: Matthew Fishman, feesh96@vt.edu Taber Fisher, taberf@vt.edu Yoseph Minasia, yosephm@vt.edu

Advisors: Tanushree Mitra, tmitra@vt.edu

Abstract: The ideal news source is one that has no inherent bias and conveys information in a straightforward manner. However, there is no such thing as a truly unbiased article. When people read news articles, they are often oblivious to the bias that the author or news organization possesses and imposes on them. Some people unknowingly read news from only one side of the political spectrum, limiting their perspective and subjecting themselves to a filter bubble. PolitiGauge aims to battle filter bubbles and enlighten the public on the bias associated with the news articles they read through a Chrome extension that utilizes machine learning to accurately gauge a given article's political bias. Several past studies relating to ours have tried to detect political bias by either taking into account bias from organizations, articles, or political speeches. What makes our approach unique is our consideration of all three components, and implementing them in a widely-accessible Chrome extension. Each component can play a vital role in the creation of our models and increase the probability of an accurate political bias detection. Another factor that makes our approach unique is moving this problem from the realm of discrete predictions to continuous predictions; the vast majority of studies found in our literature review used only three or five classes to give a prediction of bias, and the few that tried to predict on a continuous scale only used one type of data. In short, what makes PolitiGauge unique is our belief that a complex system like political bias cannot be represented by only five labels, or predicted from only one type of data.

C29: Influence Analysis of Information Operations on Twitter

Authors: Heather Robinson (hnr2800@vt.edu) Liz Dao (lizdao@vt.edu) Kayla Moore (kmoore16@vt.edu)

Advisors: Tanushree Mitra (tmitra@vt.edu)

Abstract: In January of 2019, Twitter's Elections Integrity department released the data of over 9,000 accounts associated with Information Operations in Russia, Iran, Venezuela, and Bangladesh. By analyzing each country's primary tweet topics, frequencies, and languages, we hoped to find their targets and motives. Furthermore, by comparing these findings to the same statistics of an average user from the same country, outliers stand out and speak volumes. From our findings, along with a qualitative analysis of these countries' relations, we have probable cause to assume the targets of these Twitter information operations. Russia and Venezuela may be targeting the United States, NATO members, and EU members. Iran may be targeting the United States, Pakistan, and Saudi Arabia. Lastly, Bangladesh may be targeting the opinions of its very own citizens.

C30: Analyzing the Language of Anti-Vaccination Pages

Authors: Phillip Ngo, ngophill@vt.edu Alec Helyar, alech97@vt.edu Alon Bendelac, alonb@vt.edu

Advisors: Tanushree Mitra, tmitra@vt.edu

Abstract: The anti-vaccination movement has grown and drawn more attention to vaccine misinformation. With social media growing increasingly ubiquitous, anti-vaccination pages can now disseminate information to a wider audience. It is important to understand how these pages disseminate their information and use language in order to convince individuals and parents to uphold anti-vaccination attitudes. Through a linguistic study of public pages on Facebook, we identified specific language features that distinguish anti-vaccination page content from other child and parental advisory content. We explore differences between these key features and run several classifiers to substantiate our results. Overall, we find that anti-vaccination posts are more negative, more complex, promote movement, and contain less punctuation.

C31: Python4ML

Authors: James Hopkins, jhopkins@vt.edu Brendan Sherman, brens97@vt.edu Zachery Smith, zacs97@vt.edu Eric Wynn, veric16@vt.edu

Advisors: Amirsina Torfi, amirsina.torfi@gmail.com Edward Fox, fox@vt.edu

Abstract: We understand that in the modern world, information is available at the touch of a button, and there are countless sources for many topics; these sources can suffer from numerous issues ranging from survivorship bias to an out of touch discussion level. For this reason, we set out this semester to provide a source for information on machine learning that can avoid these common pitfalls. Our project is a modular, open-source course on machine learning. It is being built under the open source for science organization and the advisement of the client, Amirsina Torfi. The initial release version of the project includes a section for basic machine learning ideas, supervised learning, and unsupervised learning. Within each section, there are 3-4 modules focused on specific methods in machine learning, including accompanying example code for users to follow along. Users are expected to follow along with the course by section and then by module, reading the documentation and executing the supplied code sample. We chose this modular approach to better guide the users as far as where to start with the course. This is based off of the assumption that users starting with a machine learning overview and the basics will likely be more satisfied with the education they gain than if they were to jump into a deep topic off the bat. Further, users can start at their level within the course, skipping over the topics they feel comfortable with and getting into what is exactly their level. The design of the course facilitates collaboration in the open source environment, keeping the course up to date and accurate. Since the course uses reStructuredText, it is both a fully functioning standalone website and a repository, publicly available for both viewing and suggesting changes.

C32: The Mayfair Project

Authors: Cristhian Benitez benitez@vt.edu, Connor Cason connerc@vt.edu, Leeciling Chea chea97@vt.edu, Owen Pilewski owen16@vt.edu, Peter Wang wpeter@vt.edu, Junwei Yao mmk9624@vt.edu

Advisors: Edward Fox fox@vt.edu

Abstract: Imagine the power of being able to summarize extensive multi-page documents and extract the most important portions. Long legal and academic documents take a long time to read and comprehend. The Mayfair Project is a new startup company founded by two Virginia Tech alumni seeking to make this summarization technology available for use by anyone. Mayfair has funded research by the Virginia Tech Computer Science department for the summarization of legal depositions and are now commercializing the results. Our capstone team has developed three websites for Mayfair to grow their business and attract new clients. The Mayfair Project is the primary site that will act as a welcome portal to introduce users to the business. The web design was created with the intent of being both inviting and professional while highlighting important aspects of the company in order to leave a lasting first impression on clients. The second of the websites we developed is called Betterdep. This site was implemented specifically for legal deposition summarization and allows users to create accounts to upload and store files. In addition, users can save summarized documents to their accounts and share any of these files with other users. While Betterdep focuses on legal depositions, there is no reason this powerful summarization technology should be limited to the legal community. Esumry, the third and final site, allows users to summarize any other document they would like. From Wikipedia pages to Terms and Conditions of software, Esumry allows users to experiment with the summarization technology Mayfair has developed. Users can copy and paste large portions of text or upload text-heavy PDF documents. While the three websites serve as prototypes for the Mayfair company, they are significant strides at improving user interest and company exposure.

C33: Adult Day Services Video Assessment

Authors: Christopher Beemsterboer cebeem@vt.edu Tyler Zebina tylerz@vt.edu

Advisors: Edward Fox fox@vt.edu

Abstract: Adult Day Services is an organization at Virginia Tech that provides person-centered care to older adults who need assistance. Adult Day Services also aims to promote the physical, social, emotional, mental, and cognitive health of its participants, and they use a variety of assessments to measure overall well-being and participant progress. These assessments are conducted in the form of interviews, and the body language, tone, and speech of the interviewer are key to performing them successfully. The training video system we created covers five different types of assessments and is designed to efficiently train new instructors to conduct these interviews. We filmed an Adult Day Services instructor conducting interviews with five different participants, each completing the five assessments. We edited the footage and compiled all of the clips of each type of assessment together. We then created a menu system which allows a user to select to play all of the training videos at once, or to play just the training video for a specific type of assessment. We delivered this project in the form of a Blu-ray .iso file which contains the menu system and the associated videos.

C34: Engineering a Cardboard Database

Authors: Leeciling Chea- chea97@vt.edu Cody Hornberger- cjh2259@vt.edu Peter Nguyen- petern1@vt.edu Shuai Liu- shuail8@vt.edu Wenjia Song- wenjia7@vt.edu Zhennan Yao- yzhenn5@vt.edu

Advisors: Rich Charles- rich08@vt.edu

Abstract: Imagine that you own a hobby store, and one of the hobbies you support is Magic: The Gathering. You want to be able to purchase and sell cards at competitive prices and keep track of the cards you have in the store as well as information about the cards. Surely, you would loathe approaching this using the pen and paper method because there are over 18,000 cards that you may need to keep track of. To our knowledge, there is no available MTG database management system which allows a business to provide their customers a view of which cards they have in stock alongside real time pricing of those cards. Cardboard Database aims to solve this problem. Our system provides a simple interface which allows the employees at Fun-N-Games Hobby Shoppe to conveniently update their inventory in bulk via a CSV file generated by a mobile application that they use (TCGPlayer). In addition, our system gets real time pricing on cards from the community's most reliable pricing source, to which the employees may apply a price multiplier to reflect their desired sale prices. Furthermore, the system supports the ability to search any card by name and view its details, such as in store quantity, price, and card text.

C35: Information and Computer Technology for Indigenous Knowledge Healthcare

Authors: Victoria Todd vtodd15@vt.edu Michael Secrist mbsec13@vt.edu

Advisors: Onwubiko Agozino agozino@vt.edu Edward Fox fox@vt.edu

Abstract: In today's world, with modern, western healthcare still unavailable to a large amount of the world's population, people are turning to complementary and alternative forms of medicine. Our goal is to create a searchable website of research supported cures that are complementary to traditional medicine, but more accessible. Our database will primarily include African indigenous healthcare information, but also pull from India's and China's complementary healthcare practices. This website will be primarily aimed at those in African countries, but will be available all over the world. We want our database to be extensible, so healthcare officials as well as lay-healers will be able to make posts containing their own tested remedies, as well as any supplementary articles for those interested. These posts will be moderated to ensure no dangerous information gets through. The posts will also be interactive, encouraging comments and ratings from those who use these remedies. A privacy policy as well as legal disclaimer are present to ensure proper precautions are taken should harmful information get through. Lastly, to support further growth of the website we will include a donation link as well as a link to our client's book. The website (health.cs.vt.edu) was created using Wordpress to take advantage of its many built in systems, including login, extensibility, social, and post moderation functionality. This will also make it easy to allow the project to be passed onto new admins and moderators in the future who may not have the technical knowledge to maintain a traditionally coded website. Since not everyone has access to a desktop computer, our chosen Wordpress theme also ensures the website is mobile friendly.

C36: Gamifying Good Behavior

Authors: Lawrence Glick lawg97@vt.edu Richard Patten richp@vt.edu

Advisors: Edward Fox fox@vt.edu

Abstract: There is a misconception that gaming and societal excellence are opposing forces; A child or young adult that spends time gaming is spending time away from study and chores. Odjinn LLC's prototype Behavioral Rewards Mobile Application, Learn2Game (working name), will help shift that perception. This mobile application will allow users to gift games as rewards to others for completing tasks. Rewards will be selected by each user and made visible to anyone they are directly connected with. Those connected will be able to offer a task which, when completed, grants the user their selected reward. We see pivotal disconnections in the interactions of parents, school districts and students as motivation for our alternative solution. Students can over-prioritize games while parents may neglect to understand and nurture their children's hobbies. Children and young adults - Gamers - and their parents and authority figures - Coaches - will be able to use our app to help alleviate the conflict around balancing perceived good and bad behavior relating to video games. Over time, this will work to create dialog where the Gamers see Quests as opportunities for fun rewards, and Coaches see Quests as opportunities to instill good behavior. This is an independent project whose intellectual property is owned by Odjinn LLC, and no Virginia Tech resources were used in the creation of this project or its materials.

C37: Workflow For Developing Gestures for Virtual Reality Using Machine Learning

Authors: David Thames davidct@vt.edu Noah Miller noahm@vt.edu Quan Nguyen ntq2503@vt.edu

Advisors: Mike Horning mhorning@vt.edu Doug Bowman dbowman@vt.edu

Abstract: We propose a unique workflow for developing immersive virtual reality (VR) gestures, driven by machine learning. VR is becoming a more and more common technology, and interaction techniques need to evolve to fit this new medium. However, there is a lot of complexity in creating gestures with this many degrees of freedom while still maintaining the natural immersive feel that VR promises. Our system uses “Wizard of Oz” user studies combined with machine learning to crowd source the gesture definitions to match users natural intuitions. Users are put into a VR experience asked to perform an action in the experience, and based on the motions they make, a model is trained to treat those motions as a gesture for that action. Using this technique, we anticipate gestures will be designed to match users natural intuition while reducing the work the interaction designers have to do. The key research questions we seek to address are: Can high-quality gesture training data be obtained through simple gesture elicitation study data, can machine learning be used to define more intuitive gestures than traditional hard-coded gestures, and how does our machine learning gesture recognition system increase user satisfaction and immersion in the environment. Our preliminary results suggest a high testing accuracy for gesture recognition; however, more results are necessary to analyze how this impacts user immersion and satisfaction.

C38: Tobacco Testimonies

Authors: Michael DiFrancisco miked24@vt.edu Nick Onofrio nicko5@vt.edu Nick Sorkin sorkinna@vt.edu Devin Venetsanos devinv@vt.edu Campbell Johnson campbel1@vt.edu

Advisors: Edward Fox fox@vt.edu David Townsend dtown@vt.edu

Abstract: Tobacco companies have had some of the best marketing strategies over the past century. It is well documented and known that tobacco produces both mental and physical health issues and yet these companies have found ways to remain as one of the largest businesses. The goal of our project is to assist Dr. Townsend in his research to understand Big Tobacco's strategies. This will be done by taking as many of the fourteen million documents released by tobacco companies online and presenting the data in a meaningful way so it can be analyzed. This project will be hosted on a Virtual Machine provided to the team by Dr. Fox and the VT Computer Science department. The idea is to take this documents using an API to scrape them and then present them to a Doc2Vec machine learning model. Then the model will be trained and be able to cluster the data so that it can then be analyzed and be used by Dr. Townsend to try and further his research. The final product will be well documented for how to use the system and maintain it. This way Dr. Townsend can do what he wants with system and any future developers for it can understand how the system works. This system is comprised of different online components such as a gensim doc2vec model and a fast approximate nearest neighbor similarity package from gensim to do the clustering of the data.

C39: Rentian Scaling in Brain Networks

Authors: Clark Cucinell clark96@vt.edu Lloyd Achey lloyd1@vt.edu Nicholas Benjamin nibe1114@vt.edu

Advisors: T.M. Murali tmurali@vt.edu

Abstract: The human brain has often been compared to a computational machine. We investigated the properties of multiple brain connectomes by evaluating them with Rent's rule, a concept originally used to determine the design optimality of very-large-scale-integrated (VLSI) circuits. The exponent describes a power-law relationship between the number of nodes in a network partition and the number of edges incident to that partition. A larger exponent correlates with a higher dimensionality of communication between the partitions, which for VLSI circuits means better computational capabilities. Using Metis software, we partitioned topological connectomes to calculate their Rentian exponents. We additionally calculated the Rentian exponent of a physical connectome using random 3D voxel generation. We replicated the linear relationship of the *C. elegans* network described in one of our core source papers, then calculated the exponent of two additional networks: a mouse connectivity network and a macaque cortical network. Both of these networks exhibited a similar topological rentian exponent to what we found in the *C. elegans* network. Our source paper had produced a rentian exponent for the *C. Elegans* network that slightly differed from our value, but the log-log relationship persisted.

C40: Trail Tweets

Authors: Max Adler madler12@vt.edu, Anthony Medovar tonym301@vt.edu, Aarjab Goudel aarjab@vt.edu

Advisors: Shuo Niu shuoniu@vt.edu, Scott McCrickard, mccricks@cs.vt.edu

Abstract: Twitter has become a hub for crowd-sourced data. The application provides an outlet for users to say or post anything they want. One of the major uses of Twitter is to share photos with other users. Starting with a years worth of tweets from 2017 to 2018 related to the Appalachian Trail, our team recognized that there are far too many pictures to easily navigate within Twitter itself. We created a way to view pictures related to the Appalachian Trail using an interactive tool designed to help users plan their trips and find new areas to explore. It allows for filtering by various tags, as well as sorting by different standard metrics. The web application also offers statistical data based on the information fed into the database. This has the overall goal of creating a better user experience when viewing pictures from the Appalachian Trail.