73rd Annual Fair

Connecticut
Science &
Engineering
Fair

March 8 - 20, 2021

Student Abstracts

Computer Science

Fair Categories

	Life Sciences	Physical Sciences
7th & 8th Grade Team	LT (1001 – 1999)	PT (4001 – 4999)
7th Grade	L7 (2001 – 2499)	P7 (5001 – 5499)
8th Grade	L8 (2501 – 2999)	P8 (5501 – 5999)
High School	LS (3001 – 3499)	PS (6001 – 6499)
High School Team	LST (3501 – 3999)	PST (6501 – 6999)

Special Categories

AT = Applied Technology	EE = Engineering: Electrical & Mechanical
AS = Animal Science	ET = Energy & Transportation
BE = Behavioral & Social Sciences	EV = Environmental Analysis
BI = Biochemistry	EM = Environmental Management
CB = Cellular & Molecular Biology	MA = Mathematical Sciences
CH = Chemistry	ME = Medicine & Health Sciences
CS = Computer Science	MI = Microbiology
EA = Earth Science	PH = Physics & Astronomy
EN = Engineering: Materials & Bioengineering	PS = Plant Science

Special Category Composites

Biotechnology	AS, BI, CB, EN, ME, MI, PS
Environmental	EV, EM
Engineering	EN, EE
Sustainability	EA, EN, EE, ET, EV, EM

CSEF Official Abstract and Certification

Word Count
216

Fair Category
LS

Project Number
3005

Title: Application of K-Means and Hierarchical Agglomerative Machine Learning Algorithms to Cluster Wolbachia Genomes based on Host Organism's Phylum

Student Name(s): S. Lee

Abstract:

Wolbachia, a species of endosymbiotic intracellular bacteria, is most notable for its role in vector control strategies to reduce the spread of diseases that are transmitted by mosquitoes, such as malaria and dengue fever. Despite this significance in disease prevention and decades of research, the exact mechanism of Wolbachia's parasitism at the genomic level has yet to be fully understood. Searching for patterns in Wolbachia's genome sequence with machine learning algorithms may offer insight into this mechanism. This project utilizes both the k-means and hierarchical agglomerative clustering machine learning algorithms to explore how Wolbachia's genome sequence is influenced by its host organism's phylum. From the NCBI database, Wolbachia genomes from both arthropod and nematode hosts were collected then preprocessed. Cluster analysis was performed on this data and visualized using Voronoi diagrams and dendrograms, which both showed stark clusters of genomes organized by the host organism's phylum with a high accuracy of 88.5%. These findings indicate a strong correlation between Wolbachia genomes and its host, suggesting that Wolbachia species adapt their genetic sequence based on its host organism's phylum. This result is promising in enhancing the understanding of Wolbachia's parasitic mechanism and also offers potential future studies, such as searching for the specific genes that differ for Wolbachia in the various hosts and applying classification algorithms.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)
MI CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Genes Dysregulated in Obesity Increase the Likelihood of Breast Cancer Development

Student Name(s): J. Boyar

Abstract:

The American Cancer Prevention Study II found that mortality due to all cancers was 52% higher in obese men and 62% higher in obese women, compared to individuals with a normal body-mass index (BMI). Among all cancers, breast cancer is the most commonly diagnosed and affects 12% of women in the United States alone. In this study, we sought to determine if gene dysregulation is a factor that increases the likelihood of breast cancer development in overweight and obese women. To achieve this goal, gene expression data from 300 women were analyzed using RNA-seq. Weighted Gene Correlation Network Analysis (WGCNA) was then used to determine the modules of coregulated genes among the cohorts. These modules were compared to specific variables to determine the relationships between the modules and traits. Finally, the pathways of the most significantly dysregulated modules were analyzed to determine which bodily functions they impacted. The WGCNA analysis resulted in a single module that was highly upregulated with BMI and contained genes associated with various immune, collagen, and extracellular matrix pathways. This module included the gene TGFβ1 - a gene that has been repeatedly linked to cancer development - thus validating the results of our studies. The module of genes upregulated with BMI is likely linked to breast cancer development in obese women. In the future, a method of inhibiting these genes would decrease the likelihood of breast cancer development in not only obese women, but also in others with similar breast cell makeups.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Word Count
250

Fair Category
LS

Project Number
3011

Title: Spider Silk Amino Acid Composition and Its Impact on Structure Function: A Novel Computational Approach

Student Name(s): H. Servin-DeMarrais

Abstract:

This project analyzed how the functionality of spider silk is impacted by its amino acid composition. All spider silks are composed of the protein family known as spidroins. There are a variety of spidroins that produce each silk type. Silk properties like flexibility and tensile strength are owed to the protein's amino acid composition.

The hypothesis stated that there is a correlation between different silk types with similar capabilities and their most abundant amino acids. Amino acid composition of spidroins greatly impacts the function of silks, meaning that amino acid makeup is a determinant of structure function.

To test this theory, I coded a tool using Python that displays a breakdown of the amino acid composition per sequence. It was determined that the spidroins that possessed similar functions to each other were the Tubuliform and Aciniform spidroins, and the Major and Minor Ampullate spidroins. Tubiliform and Aciniform spidroins are both used for wrapping mechanisms, while the minor and major ampullate spidroins are associated with web building. Analysis of the spidroins using the tool displayed a proportional makeup of amino acids seen in the genome of each pair. The Tubuliform and Aciniform spidroins had the same top five most abundant amino acids. The Major and Minor Ampullate spidroins also shared matching top five amino acids.

It was determined that amino acid abundance impacts the function of the silk it produces. Through this data, spider silk genetics can be better understood and replicated in the future for medical, engineering, and industry purposes.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) CB CS BI

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Word Count
230

Fair Category
LS

Project Number
3022

Title: Use of Accelerometers to Determine Gait Irregularity

Student Name(s): V. Calmon Coelho

Abstract:

Examining a patient's gait can give more insight into the patient's condition. One method to assess a gait is through accelerometers. Gait assessments can be done with a motion capture camera, but a less invasive approach, such as accelerometers, would allow for broader use. The objective of this study is to differentiate gait irregularity through the use of an accelerometer. It is predicted that variations in time between successive steps will identify gait irregularities. A rectangular area will be set up with accelerometers lined up on one long side of the rectangle. A subject with no conditions will walk around the square simulating five gait irregularities. Another trial will be run with a natural gait as a control. The accelerometers will record ground force data. The ground force data will be analyzed in MATLAB to remove background noise and manipulate the data. The data showed a slight difference between a normal gait's time between successive steps and that of a simulated gait irregularity. This will allow accelerometers to be used in a wider range of situations, such as home use. Home use gives more data, resulting in a better gait assessment. With more knowledge about gaits, medical professionals will be able to give better care, such as a physical therapist giving more specific treatments. Additionally, eventually, accelerometers can be used to create a wearable gait assessor, allowing for widespread use.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)
ME CS EE

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☐ Yes ☒ No

3. This project was conducted at a Registered Research Institution. ☒ Yes ☐ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Word Count

251

Fair Category

LS

Project Number

3071

Title: The Roles of Gene Expression of Non-Homologous End Joining in the Progression of Ovarian Cancer

Student Name(s): E. Lavi

Abstract:

DNA damage is repaired by the cell in multiple ways. When the main pathway is unavailable, the cell resorts to its backup, Non-Homologous End Joining (NHEJ). Causing a high level of mutations, NHEJ can both lead to cell death or the development of resistance to therapies, which decreases the chance of a patient's survival. It was hypothesized that NHEJ gene expression can predict prognosis in ovarian cancer patients. A literature search identified 48 genes involved in NHEJ. Data from 651 ovarian and breast cancer patients were extracted from the Cancer Genome Atlas. A neural network was programmed to analyze complex data patterns, with the goal of predicting prognosis. The model was trained on 332 ovarian and 274 breast cancer patients. The breast cancer patients were added to augment its learning. The trained model was then tested on a separate set of 45 ovarian cancer patients. A feature importance analysis was conducted to determine the genes most important to the achieved accuracy. Compared to random predictions, the machine learning approach showed 22% higher relative accuracy. In the feature importance analysis, genes EXO1 and TP53BP1 of the 48 stood out as most important to the predictions. These findings support the hypothesis that the NHEJ gene expression in ovarian cancer is a possible biomarker of tumor mortality. The accuracy of the model could be increased by including other variables such as demographics or tumor stage. Further inquiry is required to evaluate the potential clinical applications of the two genes identified as most important.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

ME CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☐ Yes   ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Title: Using Deep Learning to Efficiently Quantify the Impact of Road Salt Runoff on the Swim Performance of Wood Frog Tadpoles

Student Name(s): M. Zhang

Abstract:

In recent decades, roads have undergone a staggering expansion upward of 60 million km in length for the U.S. alone. These conduits of salt pollution permeate freshwater habitats, leading to rising extinction rates of 4% annually. Local wildlife like wood frogs, a widespread species vital to ecological diversity, have been severely impacted by salinization. Yet, there is relatively little known about tadpoles' reaction to de-icing compared to adults. Thus, this study focuses on how road salt exposure influences wood frog tadpole swim performance, pivotal to predator evasion. Due to restrictions, my supervisor exposed 11 groups of tadpoles to four salt concentrations up to 7,000 µS/cm and administered burst-swim trials. 650 videos overall were trimmed and analyzed using deep neural networks, a set of algorithms based on the brain capable of extremely efficient motion analysis, from the Python package DeepLabCut. Tadpole tails in 200 derived images of the training dataset were manually labeled and the network underwent 13,500 iterations to learn tail-tracking. The network then analyzed videos autonomously, providing tail coordinates for each frame. The R package trajr was used to analyze these coordinates for sinuosity, an angular measure of motion flexibility that better characterizes tail fluctuations and fitness. Results suggest that although swimming ability decreased as road salt first increased, a potentially adaptive natural safeguard was triggered by harsh conditions since the greatest sinuosity values occurred at the highest salt level. Consequently, this encourages novel conservation strategies, working in tandem with evolutionary processes, to mitigate contamination.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   EV BI CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Virtually Screen Drug Candidates to Target SARS-CoV-2

Student Name(s): L. Audie

Abstract:

SARS-CoV-2 is a virus that targets lung cells which could cause severe flu-like symptoms. It is an acronym for severe acute respiratory syndrome- coronavirus-2 (COVID-19). As of right now there are simply no cures for this potentially fatal disease. The goal of this experiment was to use the computational drug software YASARA, to virtually screen a database of natural, polyphenol molecules that would prevent the spike protein RBD from binding with the ACE-2 receptor. Particularly, Polyphenol Flavanones and Isoflavanoids were used because they were the most bioavailable and orally bioavailable of the Polyphenols. In order to accomplish this experiment, the contact residues 500, 501, and 502 were identified to use as a target for virtual screening. The top ten polyphenol molecules with the highest predicted binding affinities were selected to be docked in order to produce more accurate results. It was concluded that Naringin 6'- Malonate, Naringin, and Neoeriocitrin were the best drug candidates because they met a series of criteria: they had the highest binding affinities, they were in contact with resides 501, 502, or 503, and they formed at least one hydrogen bond.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

BI CB CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Word Count

250

Fair Category

LST

Project Number

3504

Title: Development of a WatchOS Application that uses HRV and Vibrational Pulses to Reduce Stress

Student Name(s): S. Munim, A. Liu

Abstract:

Stress is a reaction that can negatively affect many people, and it has a strong correlation to Heart Rate Variability (HRV), a measure of the variation in time between each heartbeat. Research shows that vibrational pulses can reduce stress and increase HRV. The application that was created helps reduce this stress and increase HRV for people that are negatively affected. To test the application, participants were split into 2 groups, A and B, who remotely completed a stressful task of summarizing an article in the time span of 10 minutes. During the task, Group A did not revive the vibrations whereas Group B did. Before and after the task, a VAS Questionnaire was given to measure anxiety and stress levels. HRV was also measured using the application during the task. Analyzing the VAS Questionnaire, HRV averages, and task scores, the preliminary data suggests that Group B, which had received the vibrational pulses had a higher HRV and task score, while maintaining a lower result on the VAS Questionnaire. A lower score on the VAS Questionnaire signifies less stress. Greater implications of this application are to add features such as custom haptic responses, night time vibrations to enhance REM sleep, and to send vibrations when it detects irregular HRV. A separate device can be created using HRV and vibrational pulses to make the technology more affordable for people without apple watches. These implications intend to increase HRV and reduce stress, with the aspiration of improving wellbeing for people across the world.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   BE CS AT

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☒ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No
4. Is this project a continuation? ☒ Yes   ☐ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Word Count
252

Fair Category
P7

Project Number
5002

Title: A Covid-19 Social Distancing Device Using a PIR Motion Sensor and a US-100 Ultrasonic Distance Sensor

Student Name(s): S. Srinivasan

Abstract:

Social distancing has become a huge part of our lives in the past year. It is critical we keep ourselves safe to prevent the number of Covid cases from increasing further. My device assists in doing so by calculating how close the user is to another person nearby. This device comprises of two major components: a Passive Infrared (PIR) motion sensor and a US-100 ultrasonic distance sensor. This product allows the user to customize the distance in feet they want to be socially distanced, and can also arm or disarm the system as needed. When the device is armed, the motion sensor checks for movement within its range. Once motion is detected, the US-100 calculates the distance between the user and the person nearby. If this value is greater than the distance configured by the user, an RGB LED turns green, indicating that the user is in a safe zone. If not, the LED turns red and a text alert is sent, informing the user how close he/she is from the person nearby. As a test for accuracy, I used a measuring tape to find the actual distance between the device and a moving object and compared it to the distance calculated by the device. I performed this experiment for ten days and recorded the results. Overall, my device is fairly accurate and is only 2-3 inches off at most. In the future, I plan on creating a simple mobile app to make user interaction with my device easier.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   CS AT

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Remote Learning on a Budget

Student Name(s): J. Benin

Abstract:

During the COVID-19 Pandemic, online learning has become a key factor in our education. Some families have had trouble supplying their kids with the resources they need to learn virtually. Thanks to my research and testing there is a way to not spend a fortune on a computer, but still be able to learn! A Raspberry Pi is a low-end computer that is very inexpensive. I bought a Raspberry Pi 4 and tested the things you would do when you learn online. It could easily complete the tasks. I measured the difference between a high-end Windows 10 desktop and the Pi. The Raspberry Pi was only a few seconds slower than the Windows 10 computer and it was able to perform all of the tasks needed for middle school virtual learning. While the Pi is also able to do some more advanced things like programming and 3D modeling, these tasks will go considerably slower. The Pi runs on a Linux operating system, so some games aren't supported (such as Microsoft Flight Simulator 2020).

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

AT CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Word Count		Fair Category	Project Number
233		PS	6004

Title: Bird Identification Software Using Python and Keras CNN

Student Name(s): l. lapierre

Abstract:

Bird watching is a popular hobby among many but being able to identify birds out in the field or from a picture can be challenging for those with less experience. The goal of this project is to design a bird identification software using a deep learning model called a convolution neural network or CNN. This software will allow the user to input a photo of a bird and have the software output the name of that specific bird. The convolution neural network will be designed in Python along with the Keras deep learning library. A dataset of images made by Caltech which has images of 200 bird species will be used to train, evaluate, and test the software. All images were resized to have the same dimension before passing them through the model. The model designed for this project has three convolution layers, one hidden layer and one output layer. After the validation images were passed through the network a final accuracy of 53% was reached. The test accuracy reached 98%. Further iterations of the model were done to improve accuracy and the highest validation accuracy achieved was 61%. These results indicate that the model overfit the training data causing it to do not as well on the validation images. In the future, transfer learning will be explored as an alternative way of making a CNN as opposed to training one from scratch.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Word Count
249

Fair Category
PS

Project Number
6008

Title: Using Fractal Metamaterials with Non-Integer Dimensionality to Manipulate the Propagation of Acoustic Waves

Student Name(s): N. Shell

Abstract:

A "metamaterial" is an engineered structure whose properties are determined by its macroscopic geometry, rather than its microscopic constituent materials. Metamaterials are used to manipulate waves – light, sound, or even seismic waves – in ways that may be impossible in naturally-occurring materials. However, existing metamaterials have limited bandwidth – they function only within a small frequency range, determined by their fixed geometric length scales. I simulated "fractal" metamaterials, to investigate whether the cascading length scales of a fractal can increase the functional bandwidth. Specifically, I investigated whether a Sierpiński triangle – a set of self-similar triangles of ever-decreasing lengths – can provide acoustic isolation over a larger frequency range than a simple triangle. [I optimized the acoustic band gaps – frequency ranges in which sound cannot propagate due to the structure of the metamaterial.] I used the finite-element analysis software COMSOL to simulate Sierpiński triangles of orders n = 2, 3, and 4, extruded from a 2D plane and built in 3D space. I simulated a 5-cm structure made of the plastic PLA (polylactic acid), which could be 3D-printed at low cost. I calculated the amplitude of sound transmission through the Sierpiński metamaterials over a range of frequencies from 0-20 kHz, and found a shift of the transmission spectrum to higher frequencies as the fractal order increased, as would be expected from the decreasing length scales, available to trap higher frequency sound. These expanded frequency scales may be useful for soundproofing applications, enabling the operation of sensitive scientific instruments.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)
EN AT CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: An Algorithm for Object Detection and Avoidance in Drones

Student Name(s): W. Wei

Abstract:

In recent years, the development of UAV (Unmanned Aerial Vehicle) or drones are becoming more popular with more applications including in military, industry, and entertainment. Drones can also be used in Drone swarming which is a system where a swarm of drones communicate with each other in flight and make decisions autonomously as a single unit; however, one large difficulty for drone swarms is object detection and avoiding the object in an organized manner. The purpose of this project was to create an effective method in order to determine and orderly avoid an object ahead of the drone swarm so drone swarms can navigate through more situations and locations than just in open air. The minimum requirements for this project was for the drones to be able to identify objects and coordinate paths around the object in simulations. The limitations to this project were the amount of drones so the scalability can not be tested. The solution used a Conflict based search algorithm. In order to complete this project, a drone simulation program was needed. The drones and the research platform is at Sacred Heart University's MakerSpace and they were used under the supervision and guidance of Professor Tolga. It was found that the solution was effective in directing the drone swarm to the target location With this algorithm, drone swarms can go into more terrains and locations to reach its goal, which will increase the areas which the drone swarms can be useful.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☒ Yes   ☐ No

4. Is this project a continuation? ☒ Yes   ☐ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Word Count
236

Fair Category
PS

Project Number
6012

Title: Modular Robotic System for Multipurpose Robots

Student Name(s): B. Ochs

Abstract:

Currently there are robotic systems that replace human first responders, but they have only one specific use. This project will design a robotic system that will demonstrate the ability to replace first responders in a dangerous situation, no matter what the situation is by being modular. A possible solution to a modular robot will be designed, developed, and a small scale, 3D printed prototype will be built. It would use very similar electronics to a proper, fully-functional robot, but it will only have basic functionality as its primary goal is to demonstrate modularity. To connect the modules of the robot there will be dedicated high power plugs and a primary plug will carry all data and lower voltages like 3.3V and 5V. To prove that the modularity works properly, multiple modules will be made and be tested for their ability to connect to the robot and interchange with each other. The current basic tests have given promising results. A central computer is able to communicate with external microcontrollers simulating a module. The programming of more simulated modules is planned and construction of housings for the computers is in development. The method of communications used is favorable for these applications, as it is easy to use and only requires two data lines. The concept of using a standard connector to carry these signals has also been tested and should not cause any problems later in development.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)
EE AT CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☒ Yes   ☐ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Title: Incorporating Pre-Game Rankings to Colley, Massey, and Elo Methods

Student Name(s): J. Liu

Abstract:

Colley, Massey, and Elo are three popular rating methods commonly used in the NFL and Major Division I NCAA sports where game results from the regular seasons can be used to seed major playoff brackets or create predictive rankings. As with most rating systems, an ample amount of game data is needed for the resulting rankings to be reflective of the competitors' strength relative to others. However, in situations such as a professional golf match play tournament, where there is a lack of available match play data, or initial weeks into regular NFL season, when it is too early to have the appropriate amount of needed data in hand, how can we produce rankings that meet the standard of accuracy? To address this problem, this project proposes two possible seeding methods, the dominance graph method and the weak dominance graph method, that both incorporate pre-game rankings to make up for the lack of relevant games that are adaptive to Colley, Massey, and Elo. In predicting the outcomes of the Dell Technologies Match Play and weekly NFL games outcomes, the effectiveness of the method can be tested.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   MA CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Word Count

250

CSEF Official Abstract and Certification

Fair Category

PS

Project Number

6027

Title: Project Regrow: Cutting-Edge Technology to Fight Deforestation

Student Name(s): S. Bhardwaj

Abstract:

In the 2020 wildfire season alone, the California Department of Forestry and Fire Protection projected 4,197,628 total acres of land were burned throughout California, ravaging countless forests and wildlife. These devastating wildfires have been drastically affecting wildlife for generations and have been adding to the fast-growing issue of climate change. Project Regrow intends to alleviate this fast-growing problem. Project Regrow intends to protect, preserve, and grow the world's forests by using drone technology to plant seeds. The drone itself can be filled with seeds, and flown over the desired area for regrowth.

Project Regrow utilizes a 3D printed skeleton, housing the Raspberry Pi Zero W, a battery pack, a seed container, and a seed disperser to disseminate seeds. The 3D printed skeleton can be mounted onto any drone that is at least 36 millimeters wide and 31 millimeters tall with the use of Velcro straps to secure it. The seed disperser opens and closes to drop the seeds from the attached container by using a servo motor. The servo motor is connected to the Raspberry Pi Zero W, which is programmed in Python to open and close the shutter every 2 seconds, which gives enough time for the seeds to disperse. This process also saves enough seeds in the container to maximize the area for the seeds to be planted. To prevent excess seeds from being dispersed, an adapter between the seed disperser and the seed container was made. More development is necessary to ensure full forest revitalization.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) AT EE CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Artificial Intelligence Medical Question-answering Robot

Student Name(s): Y. Qi

Abstract:

With the development of the medical system over years, the problem of wasting much time in a line for only a simple diagnosis or match medicine is still unsolved. However, with the development of artificial intelligence, automatically answer questions based on Deep Learning can be accomplished and thus being able to solve the problem. The experiment focuses on the accuracy and loss of the model based on the variety and amount of data under the same epoch size. To process human language, NLP (Natural Language Processing), a field focusing on human language and computer science, can find the keyword among datasets and thus learn therapy from the data about different illnesses and doctor-suggested treatments. By using Transformer model as a basic frame, features of the content of each QA pairs inside the dataset are remembered and the relationships among them are calculated and record. Through the model, the robot can output the match advice under specific conditions based on the questions asked by patients. After experiments on comparing accuracy with different amounts and variety of data, the conclusion confirms that a larger dataset with higher variety provides higher accuracy for the output when the number of training epochs stays the same. Among the five datasets, when the number of epochs is two hundred in the training, the model based on the largest dataset achieves the optimum efficiency - the accuracy is 0.971 while loss is 0.004.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) CS ME

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Word Count
246

Fair Category
PS

Project Number
6050

Title: System for Underwater Passive Identification of Acoustic Signatures of the Delphinapterus leucas

Student Name(s): I. Mendiratta, I. Mendiratta

Abstract:

Though there are pre-existing programs utilizing the Fourier Transform to recognize specific vocalizations of bats and aves, there is yet to be such a successful underwater contraption. For species like the Stenella Clymene, a like system which utilizes deep learning to identify and track dolphins based on signature whistle types (SWTs) can provide a currently unknown estimation for the population. Due to limited access to the S. Clymene, initial programming was done based on vocalizations of the Delphinapterus leucas (Beluga whale), acquired from a local marine life reserve. These audios were inputted into a neural net to be broken down by the discrete fourier transform into their wavelengths such that they could be used as the control wavelengths searched for in unclassified audio. The program works by using the Transform to separate audio into its individual wavelengths and then by isolating those deemed important, that is part of a possible S. Clymene audio signature. The goal of developing an AI program equipped to detect specific vocalizations from mixed audio is completed at a rudimentary level, as a foreign D. Leucas audio can be detected based on its similarities to a known and inputted audio - as identified by a neural net. The AI program created can be adapted to account for relative variance in S. Clymene vocalizations. The program can ultimately be applied to provide a new way of monitoring population analysis of marine life, in a larger scheme, specifically including that of the S. Clymene.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)
CS AT EV

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No
3. This project was conducted at a Registered Research Institution. ☒ Yes   ☐ No
4. Is this project a continuation? ☒ Yes   ☐ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Word Count		Fair Category	Project Number
251		**PS**	**6052**

Title: Creating a distributed denial-of-service attack mitigation system using SYN cookies to ensure access to crucial election information

Student Name(s): J. Krukar

Abstract:

The need identified for my project is a strong system to protect against cyber attacks, specifically TCP SYN flood attacks, on websites that contain important information for the elections. Around election time, the election may be influenced by data breaches and a lack of accessibility to important information. The most common type of attack is a DDoS attack, a malicious attempt to disrupt the normal traffic of a targeted server by overwhelming it with a flood of Internet traffic. Use of an Anycast DNS ensures that any one of a number of DNS servers can respond to DNS queries, which reduces latency and improves uptime for the DNS resolving service. The criteria of my project is a specific focus on election data protection. Constraints are a short timeline and limited resources. My solution to the identified need was to build a mitigation system using Anycast DNS and SYN cookies to carry out mitigation, which allows the server to quickly filter through the queue. The mitigation system was programmed in C++ then tested against an open-source DDoS attack in Ubuntu Linux employed by Kali Linux. The strength of the mitigation system was analyzed based on the collection of data on how fast it detected and mitigated the DDoS threat. The data suggested that the designed system was stronger and faster than the basic DDoS mitigation system. The system has the potential to be implemented throughout small-server websites throughout the entirety of the Internet to better protect them from DDoS attacks.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Designing a Ground Collision Avoidance System to Inhibit Controlled Flight Into Terrain Accidents For Single Engine Aircraft

Student Name(s): M. Beaudette

Abstract:

Controlled flight into terrain (CFIT) is an accident in which an airworthy aircraft, under pilot control, is unintentionally flown into the ground, a mountain, or an obstacle. CFIT accidents were identified as a cause of 25% of USAF accidents. The main goal of this project is to design a ground collision avoidance system (GCAS) that will identify CFIT accidents before they occur. Many articles regarding CFIT and GCAS systems highlighted that these did not appear in small aircraft. The GCAS software was flashed onto an Arduino. To identify a CFIT accident is impending the code uses Airspeed (knots), altitude, and pitch. These values were supplied by the accelerometer, to determine airspeed, the barometer/GPS to determine altitude, and the gyroscope, to determine pitch. The data produced was fed into an algorithm that calculated the time before collision with the ground. If the time exceeded a threshold, the code sounded an alarm. In addition, the code used map data to determine the highest points. From that the code took a radius of highest points and compared it to the aircraft's altitude. If the code identified that the aircraft's altitude conflicted with a high point, the pilot was told to increase altitude. As proof of concept, the system was placed into an enclosure that can fit in a cockpit. The outcome is to provide a small, affordable system that will decrease CFIT accidents and can be retrofitted by pilots.

**Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)** CS EE AT

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☒ Yes   ☐ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Title: Creating an Algorithm that Automatically Scores Tissue Images Like a Pathologist

Student Name(s): I. Yan

Abstract:

Cancer is the second leading cause of death globally. Detecting and treating cancer early on can save many lives. Pathologists have to look at tissue microarray (TMA) images manually to identify tumors, which can be time-consuming and inconsistent. Existing algorithms that automatically detect tumors have not achieved the accuracy of a pathologist so far, so they are not used widely.

A major challenge is that TMA images with different shapes, sizes, and locations can have the same score. Learning staining patterns in TMA images requires a huge number of images, which are severely limited due to privacy concerns and regulations. TMA images from different cancer types have common characteristics that could provide valuable information, but using them directly harms the accuracy.

For the first time, transfer learning was used to enlarge the training sample by extracting prior information from tissue images of different cancer types. The cancer TMA images were scored based on the severity of tumors. Images used in developing the algorithm were taken from the Stanford Tissue Microarray Database. The accuracy was calculated by comparing the score given by the algorithm to that of pathologists.

Transfer learning has made it possible for the algorithm to break the critical accuracy barrier. Pathologists had 75% accuracy, while the algorithm achieved 75.9% accuracy. This will allow pathologists to confidently use automatic algorithms to assist them in recognizing tumors consistently with a higher accuracy in real time.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) AT CS ME

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents
☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Word Count
258

Fair Category
PS

Project Number
6062

Title: Meta Programming in JavaScript

Student Name(s): J. Colangelo

Abstract:

My goal was to research Meta-Programming and try to create Meta-Programs using Javascript. My research showed that a clean meta script, a program that generates code without outside influence, wasn't possible using Javascript. However, a non-isolated program allowed for creating a script that was similar to meta-programming in Javascript. This recreation is essentially a large JavaScript library. The first part of these functions are dedicated to If protocols and Array protocols, which run complex if statements, including nested if statements, in a single line. The Array Protocols provide for loops that can run through arrays and statements that can declare entire arrays. The second part of the Meta Library are the math-based functions and the OPER constructs. The math-based functions allow for simple or complex calculations, processing a maximum of four equations at a time. They include prewritten functions for common algorithms, such as background scrolling. The best part of the math-based functions is the OPER constructs. These constructs are rapidly processed algorithms that use O, P, E, and R as variables. OPER can perform thousands of operations with input. I wrote the OPER constructs to be used as parameters for the math-based functions, combining the processing power to exponentially increase the speed and capabilities of the program. While a true, clean Meta-Script was not possible, the Meta Library allowed for restructuring the syntax of a program on a massive level to write simpler code that could do more things, the closest possible Javascript recreation of clean Meta Programming.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☐ Yes   ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Word Count
205

Fair Category
PS

Project Number
6063

Title: Development of an autonomously navigating robot capable of conversing and scanning body temperature to help screen for COVID-19.

Student Name(s): R. Kim

Abstract:

Throughout the COVID-19 pandemic, the most common symptom displayed by patients has been a fever, leading to the use of temperature scanning as a preemptive measure to detect and isolate potential carriers of the virus. Human employees with handheld thermometers have been used to fulfill this task, however this puts them at risk as they cannot be physically distanced when taking forehead temperature readings and the sequential nature of this method leads to great inconveniences and inefficiency. The proposed solution is an autonomously navigating robot capable of conversing and scanning people's temperature to detect fevers and help screen for COVID-19. To satisfy this objective, the robot must be able to (1) navigate autonomously, (2) detect and track people, and (3) get individuals' temperature reading and converse with them if it exceeds 38ºC / 100.4ºF. An autonomously navigating mobile robot is used with a manipulator controlled using a face tracking algorithm, and an end effector consisting of a thermal camera, smartphone, and chatbot. The goal of this project is to develop a functioning solution that performs the above tasks. In addition, technical challenges encountered and their engineering solutions will be presented, and recommendations will be made for enhancements that could be incorporated when approaching commercialization.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) EE CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☒ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Word Count
242

Fair Category
PS

Project Number
6064

Title: VISTA: An Smart IoT Device to Alert Clients Regarding Freezing Pipes, High Heating Bills, and Heat Sickness/Heatstroke.

Student Name(s): S. Srinivasan

Abstract:

My product, VISTA, is aimed at solving three critical problems: freezing pipes, high heating bills, and death due to heatstroke. The VISTA sensor system is built using Arduino and Adafruit technologies. The system provides a dashboard to configure the settings and sends mobile text alerts to users.

To address the freezing pipes situation, I researched and saw that a temperature reading of 55 degrees or above is preferable to avoid a pipe burst. Using this information, I decided to alert the user if the temperature dropped below that level; however, the users themselves can change this baseline temperature setting below which the user needs to be alerted.

To address heat exhaustion and heatstroke, I have to calculate the heat index value of the environment. Heat index is caused by a combination of the temperature and humidity. I used the Rothfusz regression formula to calculate the heat index value. The calculated value can be classified into 4 conditions (caution, extreme caution, danger, and extreme danger) based on the severity. I alert the user accordingly.

To target high heating bills, the system needed a way to detect air leaks. This led me to create a secondary sensor, identical in hardware, placed near windows, doors, and vents around the home. I also kept a primary sensor in the center of the house. Through my trials, I saw that a difference in 15 degrees was indicative of an air leak and thus alerted my user accordingly.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

EV CS ET

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☐ Yes   ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Title: Examining the Relationship Between Labor Cost and Firm Strategy Through Mixed Methods Content Analysis on Panel Data and SEC Filings

Student Name(s): A. Kabra

Abstract:

With a desire for growth and longevity, corporations have looked to analyze labor cost, a proven indicator for capital success, to recognize which components of firms ensure maximal revenue. Interpreting such indicators help make vital managerial decisions regarding production cost adjustment and workforce investment. In recent years, it has been hypothesized that firm strategy also impacts corporate success but could not be evaluated due to the data's qualitative nature. In this study, I analyzed labor cost and its variability in relation to five prominent firm strategies through the creation of a novel mixed-methods methodology using tokenization and root-word identification. I examined a cross-sectional dataset spanning over 30 years with 3300 observations, performing content analyses on qualitative SEC 10-K filings and multivariate regression analyses for quantitatively extracted frequencies. Multivariable statistical models were analyzed via R and SPSS and organizational control variables, such as capital expenditure, were selected due to their high correlation with labor cost, acting as a way to avoid multicollinearity. Results showed that innovation, human resources, and consumer-focused strategies had a strong association with labor cost while growth, restructuring, and human resources shared strong correlations with labor cost variability ($p<0.01$). Such results led to conclusions on the implementation, advantages, and drawbacks of each strategy and their magnitude of significance in relation to capital success. Firms can use this novel mixed-methods approach to draw conclusions and make critical organizational decisions referring to cost allocation and strategy identification. Future studies should include content analysis using multi-layered analytical techniques and big data.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)
MA CS AT

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Comparing Algorithms to Solve the Exact Cover Problem with Sudoku

Student Name(s): A. Pourkavoos

Abstract:

The exact cover problem is an NP-complete problem in computer science with practical applications in coordinating schedules. This project compared the efficiency of C implementations of the backtracking algorithm to solve Sudoku, a special case of the exact cover problem. Each program stored not only the board but also a 9x9x9 array containing which digits were possible for each cell. This allowed the program to backtrack if a cell had zero possible digits remaining, or to immediately fill in cells with exactly one digit remaining. Three backtracking variations were compared. The first was the naive approach, storing only the above arrays. The second variation also stored how many digits remained possible for each cell, updating the count as necessary. The third variation stored, in addition to the above, a doubly linked list of all possible digits for each cell, implementing Knuth's Dancing Links algorithm. Each variation has faster asymptotic behavior than the previous but also requires more operations per individual action, such as eliminating a digit. Each variation was implemented in two ways: single-threaded and multi-threaded, the latter using the POSIX Threads (pthreads) library, for a total of six programs. Of the three variations, the naive approach was the slowest, and Dancing Links was the fastest. The multi-threaded programs ran approximately 20 times more slowly than their single-threaded counterparts (1.49 vs 33.1 ms per puzzle on average), likely because the overhead of creating and terminating threads outweighed the benefits of running an already-efficient program on four cores vs one.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) CS MA AT

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☐ Yes   ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Title: Wildlife Recognition from Camera Trap Data using Computer Vision Algorithms

Student Name(s): Y. Dai

Abstract:

Camera trap is a common method to capture wild animals on film. It has been widely used in wildlife conservation for decades. With increasingly available vision data from camera traps, it becomes prohibitively costly to manually extract useful information from these data. Leveraging recent advances in machine learning, I adopted and fine-tuned a popular convolutional neural network (CNN) architecture called YOLOv3 to automate knowledge extraction from the camera trap data. Specifically, I trained the YOLOv3 model on Zooniverse's public large-scale camera trap dataset, which contains wildlife species information but lacks the bounding box information. In order to help the model highlight animals in the images, I manually labeled thousands of camera trap photos of wildlife like rhinos, lions, and impalas using a visual object tagging tool to prepare for the supervised learning. As a result, the deep learning model fine-tuned on these labeled training data are able to detect and recognize more wildlife species with high accuracy and confidence value compared to the pre-trained model. Meanwhile, compared to the existing work by Norouzzadeh et al. (2018) and Willi et al. (2019), my model further provides bounding boxes with predicted labels that can facilitate the wildlife identification and recognition. In this way, we can further reduce the human effort for scientists and zoologists to analyze camera trap data and possibly improve the intelligent understanding about wildlife at lower costs. Further work of this project may include real-time wildlife detection in camera trap videos and the design of websites or mobile apps.

**Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)**   EV CS EM

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☐ Yes   ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Title: Developing Tetherless Motion Tracking of Wandering Insects with an Omni-directional Servosphere Robot

Student Name(s): N. Lee

Abstract:

This project is the development of a servosphere robot capable of treadmill-style operation and tetherless motion tracking. Wandering insects are model organisms for researching motion patterns and how they are affected by certain genetic factors. However, conventional methods of motion tracking, such as tethers or markers, interfere with insects' natural behavior, yielding inaccurate results. An omni-directional sphere, or servosphere, can rotate in any direction around any axis, such as with "Hamster ball" robots. A servosphere could be used as a treadmill for an insect, counteracting and recording its movements.

The procedure follows three sections: designing and constructing a servosphere robot using servo-powered omni-wheels, deriving theoretical methods of servosphere operation, i.e., an inverse-Jacobian transformation matrix, and implementing operation into the robot using Python-based motion tracking and computer vision.

The servosphere robot was evaluated based on two criteria: consistency of operation as a treadmill, and pathing error. Consistency is the frequency of successfully moving the silhouette from an arbitrary position to the sphere's center, and pathing error is the divergence of the silhouette's motion from the ideal path, measured as the trajectory's standard deviation. The results were promising, with 100% consistency and an average error of 0.484cm across all angular positions.

In this project, a servosphere robot was successfully constructed and operated by implementing theoretical models, and the viability of the servosphere robot as a motion tracking system was demonstrated. It can be used in research with genetically edited insects to isolate the effects certain genes have on motion behavior.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   EE MA CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No

4. Is this project a continuation? ☐ Yes   ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

CSEF Official Abstract and Certification

Word Count

231

Fair Category

PS

Project Number

6078

Title: Development of a Novel Classification Model to Determine Vegetation Habitability of Exoplanet Atmospheres Via Random Forest Machine Learning Algorithms.

Student Name(s): J. Gottlieb

Abstract:

Based on a 2013 NASA report there are ~11 billion potential exoplanets present in the Milky Way Galaxy of these, ~4,000 exoplanets have been confirmed.   To further evaluate exoplanet candidates, data must first be filtered and sorted. Machine learning algorithms can streamline this processing time by providing definitive insight into which exoplanets require prioritization for further investigation. A novel application of a random forest algorithm was used for the analysis of exoplanet transit spectroscopy to determine the habitability for a given exoplanet's atmosphere. The data set used consisted of 20,206 transit spectroscopy points, 3,143 were collected from the Nasa Exoplanet Archive; the remaining 17,063 points were generated using Nasa's Planetary Spectrum Generator. Performance analytics data provided the model with Precision, Recall, Accuracy, and F-1 Scores for 0 and 1 Habitability rating as 0.68, 0.97, 0.82, 0.93, 0.74, and 0.95 respectively with an accuracy of 0.91. The model's ability to predict Habitability values of 1 with performance scores all above 0.9 indicates that the model is successful. Limited bias in the model's predictive capacity of inhospitable values was detected, as indicated by sub 0.9 scores. Through larger data set value input, it is expected that this bias will be eliminated. This novel machine learning model can be used to increase the efficiency at which exoplanet data is analyzed and interpreted, assisting in the prioritization of celestial bodies which merit further investigation.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

CS PH AT

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents

☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Sentiment Analysis on Twitter Data Regarding the 2020-21 Georgia Senate Runoff Election

Student Name(s): S. Jonnalagadda

Abstract:

In the age of social media, opinions can be shared easily and quickly. There are many platforms on which users can post highly biased and political messages, most of which display a positive or negative attitude towards a candidate. Twitter is a goldmine for such messages, and its API (Application Programming Interface) and the Tweepy library allow users to extract the messages. To get access to the API, I had to create and obtain approval for a Twitter developer account. Then, I had to make an app in the API in order to receive a set of keys, which must be concealed for account security. With those keys and the Tweepy library, the tweets can be extracted and saved as text in a pandas dataframe alongside the timestamp of the tweet. This text has to be pre-processed, meaning that words that are neutral will be removed and added to the dataframe. Raw tweets containing redundant information are disorganized. Pre-processing will help overcome those issues. After pre-processing, the tweets can be analyzed for sentiment by using the TextBlob library and saving the polarity in the dataframe. With the polarity saved, the data can be averaged. When 50% of the votes were counted, Perdue's average polarity was 0.113, and Ossoff's average polarity was 0.135. After 90% of the votes had been counted, Perdue's average was 0.092 and Ossoff's average was 0.197. My code predicted that Ossoff would win because of his higher average.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Comparing the Efficiency of Autonomous and Regular Vehicles With New Requests

Student Name(s): D. Luo

Abstract:

Autonomous vehicles will become prevalent among the public within a few years, aiding in the reduction of traffic through communication between vehicles. However, properly planning the routes of these vehicles and testing their efficiency has been recently subjected to research. This experiment compares the efficiency of autonomous and regular vehicles with the appearance of new requests with total time traveled as the measure. Using Python, a computer model records the total time a regular and autonomous vehicle needs to pick up every location and fulfill new requests. The autonomous vehicle can change its route with the appearance of the new requests while the regular vehicle must make a second trip. An 8x8 grid is used for simplicity. This experiment comparing the efficiency of autonomous cars and exploring the potential of rerouting can inspire scientists to pursue the development of a smarter car. The results will help the public understand the potential increase in the efficiency of autonomous vehicles.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)

AT MA CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Generalized Correction of Spatially and X-Z Correlated Errors through Adaptive Minimum-Weight Perfect Matching Topological Algorithms

Student Name(s): S. Florin

Abstract:

Topological quantum error correction uses the topology of surfaces like the plane or torus to correct quantum errors, helping quantum computing to be successful. However, topological quantum error correction assumes that errors occur independently, whereas real errors are often correlated. This project considers both spatial and X-Z correlation. The error model uses fixed-length random walks of errors, providing the spatial correlation. Also, chains are made up entirely of either Pauli X, Y, or Z errors, each with independent fixed probabilities. This forces the decoder to consider X-Z correlation. The decoder is a variation of the minimum-weight perfect matching (MWPM) algorithm. Instead of the weight of an edge being the taxicab distance, it is defined using a combination of functions on distance and chain syndrome overlap. The distance function combines models of the fixed length of the walk and partitions of the walk into staircase-shaped chains. The overlap function peaks at zero and complete overlap to account for Y chains. The structural changes to the algorithm reduce errors by >90% when compared to the error model and algorithm from the initial research. The new decoder accounts for a variety of parameters so that errors are more realistic, including error type distribution and the probability that the type of walk aligns with syndromes. Independent of these parameters, this revised algorithm provides at least a 30% reduction in errors as compared to the traditional MWPM algorithm. This shows this algorithm can successfully adapt to these random walk errors and improve quantum computations.

**Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)** CS PH MA

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☒ Yes ☐ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Title: The Hat and the Mouse: A Practical Hands-Free Method to Control a Personal Computer

Student Name(s): P. Noe

Abstract:

Due to their unique challenges, many quadriplegics experience difficulties when performing everyday tasks, like using the computer. This head mouse was designed to allow users access to full functionality of a PC mouse without needing to use their hands. The device was programmed using the C++ programming language, and it utilized various Arduinos as well as Bluetooth and voice control features. In order to determine the performance and functionality of the headmouse, experiments were conducted to measure battery life, voice control reliability and the degree of head tilt required to move the cursor across the computer screen. The voice control module was found to be 94% reliable, and the required degree of head tilt for full screen deflection was around 10°. Although the battery life was only 42 minutes, the device's overall performance was still very good. In being able to provide users with complete control of a computer mouse while still presenting minimal errors, the headmouse demonstrated its ability to carry out the functions for which it was designed. In the future, other features could be added to the device, including the recognition and implementation of key words and phrases, such as "today's news," which could be carried out using the Arduino keyboard.h library. Additionally, the device's power source could be changed to a 6 volt battery, as opposed to a 9 volt battery, to minimize the risk of overheating and lessen the weight of the headmouse. Use of Bluetooth Low Energy (BLE) could also extend battery life.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   AT CS EE

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No
4. Is this project a continuation? ☐ Yes ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Title: Comparison of Machine Learning Models for Short-Term Forecasting of Solar Power Generation Output: A Case Study

Student Name(s): D. Yuan

Abstract:

The Loomis Chaffee School's solar field is the largest solar power plant among any K-12 school in Connecticut, built in late 2019. Like all solar fields, its power generation is highly variable and dependent on external factors such as weather. For such photovoltaic power sources, reliable and successful integration into the larger power grid system depends upon knowledge and prediction of future power output. The application of machine learning models for solar power output forecasting has thus become popular in research and literature, replacing past approaches based on statistical or physical models. This project aims to determine the feasibility and performance of applying three machine learning algorithms, support vector machines; random forests; and k-nearest neighbors, for short term solar power output forecasting, through leveraging the rich data generated by Loomis' solar array in the first comprehensive study of that data. After a preliminary data exploration phase, data is preprocessed for elimination of error values and normalization. Hyperparameter tuning is conducted to improve model performance. Only in situ data is used, and findings indicate that the training of the three models on the limited set of data features has promising utility in power prediction, with a significant improvement over a baseline persistence method for hourly prediction and best performance by support vector machines. In the future, efforts to extend this research include adding local weather data and establishing a data monitoring platform for the student body at Loomis Chaffee for education and promotion of solar energy.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)   CS ET

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No
4. Is this project a continuation? ☐ Yes   ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No

Title: Python Security Functions

Student Name(s): S. Vakacherla

Abstract:

The purpose of this project is to see if a program written using Python can be simple enough for anyone with or without CS knowledge to use and also be effective in protecting against hackers. As the world continues to move online, cybersecurity becomes a more pressing issue. In this project, the program had different cybersecurity functions: a password checker that would say whether a password was weak or strong, a password generator that would output a sample strong password using random character generating, a URL decoder that would decode a percent-encoded URL, a base64 function that would be able to decode base64 and encrypt messages to base64, a timestamp function that would output the immediate time and date down to the second, and an API function that would - given an IP address - output where this address is based, what corporation it is associated with, and how many different security firms have flagged it as harmful. These different programs were created using the programming language Python and then tested on the IDE Pycharm. The results of the project were excellent - the program worked better than expected, and the user interface was easy to follow. Each program worked as intended, and the API program also was able to find other information such as how many other IP addresses are associated with the given one. In conclusion, this project proved that a python program could very well be effective against hackers, and easy to use at the same time.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) | CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects ☐ potentially hazardous biological agents

☐ vertebrate animals ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

Word Count

159

CSEF Official Abstract and Certification

Fair Category

PST

Project Number

6505

Title: Stock Simulator

Student Name(s): A. Patel, M. Gallo, E. Lombardo

Abstract:

The objective of our project is to allow people to practice investing in the stock market, and become more aware of the methods by which they can earn money.

To do this, we created a simulation of a stock market, using ProcessingJS, a JavaScript library, in which people can practice investing money and see the results of their investments.

The simulator includes 8 randomly generated stocks, which are unrelated to real world stocks. They have random price values, which randomly fluctuate similar to how real stock prices work.

Users are given $1M in starting currency, and can buy and sell any number of shares of any stock whenever they please with the currency they can afford to spend. The simulator keeps track of the earnings produced by each owned investment, as well as the earnings owned overall (in $ as well as %).

The simulator successfully models a real-world stock market and allows users to practice before spending real money.

**Technical Disciplines Selected by the Student
(Listed in order of relevance to the project)** CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents

☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes ☐ No

3. This project was conducted at a Registered Research Institution. ☐ Yes ☒ No

4. Is this project a continuation? ☐ Yes ☒ No

5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes ☒ No

CSEF Official Abstract and Certification

Word Count
240

Fair Category
PST

Project Number
6510

Title: Creating a Platform Capable of Conveying News Article Bias

Student Name(s): P. Scully, J. Feuerstein

Abstract:

Today, people of differing political ideologies tend to get their information from different sources and often have very different opinions of bias. Fortunately, computer science already has the potential for more objectively analyzing media bias. The purpose of this project was to take advantage of this and create a platform capable of dissecting articles in the context of bias and factual integrity, conveniently and succinctly conveying this information to a user.

The constraints of this project were limited time, lack of extensive knowledge or experience in design and application programming, and a lack of money to buy expensive design software. The criteria were that the platform accurately assesses bias and has a functional and user-friendly user interface. The application was programmed modularly, such that features were sequentially programmed and integrated upon readiness.

The application was programmed in Django, a Python-based open-source and free web framework, utilizing the React Native typescript library. Bias is assessed by comparing the text of articles to a list of biased words aggregated by researchers from Stanford. All source code is hosted in a GitHub repository, and all UI/UX design was done in Figma. All work with the mentor was in code reviews and advice -- no direct progress was made by the mentor.

If effective, this application will have the potential to aid in reducing the divisiveness of current politics and social issues, possibly leading to a more educated and civil future.

Technical Disciplines Selected by the Student
(Listed in order of relevance to the project) AT CS

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

☐ human subjects   ☐ potentially hazardous biological agents
☐ vertebrate animals   ☐ controlled substances

2. Student independently performed all procedures as outlined in this abstract. ☒ Yes   ☐ No
3. This project was conducted at a Registered Research Institution. ☐ Yes   ☒ No
4. Is this project a continuation? ☐ Yes   ☒ No
5. My display board includes photographs/visual depictions of humans (other than myself or my family):

☐ Yes   ☒ No