CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Kamran Ansari

Project Number
J0801

Project Title
Mathematical Modeling of Intracranial Pressure Response to Inversion Therapy for Dural Leak Headaches

Abstract

Objectives/Goals
The objective of this project was to create a mathematical model to predict the duration of response to administering inversion therapy to those with dural leak or low intracranial pressure (ICP) positional headaches.

Methods/Materials
When researching to see what mathematical models existed that predicted intracranial pressure (ICP), I found that they all relied on physiological models that fail to account for a phenomenon called a jugular vein collapse, a variable from the venous system where the jugular vein in the upright position contracts, on ICP. I then took existing mathematical and biological equations and modified them to account for a craniospinal fluid volume loss from a dural leak as well as pressure changes from inversion therapy. I then programmed my equation into MATLAB to produce graphs that predicted the relationship of ICP over time after inversion therapy in the setting of a dural leak.

Results
To see if my model correctly predicts ICP in a normal upright patient, I set variables in the model to reflect a normal patient's physiology. My model correctly establishes the initial ICP pressures of a normal person, in accordance with current values published in the literature. I then tested my model at an initial ICP of 1, 2, and 3mm Hg, which corresponds to a patient with a dural leak and, as a result, a very low supine ICP. Finally, I set variables in the model to reflect an abnormal patient's physiology that was then inverted to achieve an ICP of 12, and ran the model for 60 and 400 minutes. My model shows that, although the effect wore off as time progressed, as expected, inversion delivers sustained relief (in the form of an upright ICP in a normal 0 to -2 mmHg range) well past a few minutes (as predicted by convention) and upwards of 3-4 hours.

Conclusions/Discussion
My results have large implications for how dural leak patients are conventionally treated. My model shows that inversion therapy, using a simple low-cost inversion table, can provide pain relief for dural leak headaches for 3-4 hours. This runs counter to the conventional wisdom which states that any positive, ICP increasing effects of inversion therapy would dissipate within minutes. Replacing or at least augmenting conventional pain medications with inversion therapy could result in decrease in opioid use or invasive treatments for dural leak headaches, which effect 500,000 individuals in the US every year.

Summary Statement
I created a new mathematical model to predict intracranial pressure after inversion therapy for those with dural leak headaches, showing that inversion therapy results in relief of these headaches.

Help Received
None. I learned how to solve differential equations using derivative calculus videos on Khan Academy. I taught myself MATLAB using the MATLAB website and forums.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)

Sriram V. Bhimaraju

Project Number

J0802

Project Title

Low-Cost Archery Assistant with an Interface for the Visually Impaired

Abstract

Objectives/Goals

Archery coaches use verbal communication and visual demonstrations to provide feedback to archers. This needs to be available to them in real-time and as an exact science in practice time, so the archer can learn perfectly. The engineering goal of my project is to provide a low-cost archery assistant for sighted and visually impaired users that provides real-time form corrections, draw length adjustments, and bow positioning for optimal launch angle, to ensure accuracy during practice.

Methods/Materials

During bow calibration, readings for optimal form and draw lengths were taken with a spring scale and tape measure and mapped to the readings of the flex sensors on the bow limbs sent by Arduino. Hooke's law graph was drawn, Spring's constant measured, potential energy, kinetic energy and arrow velocity calculated. During game time, the wind speed was obtained using an anemometer, drag on the arrow computed and Euler's rule used to calculate the final travel of the arrow. The App then computed the optimal launch angle to hit the target board, correlated the real-time flex sensor values and phone accelerometer values against values noted during bow calibration, to give audible and visual corrections during shooting by an archer for better aiming at the target.

Results

I evaluated my app with a toy, medium and professional bows. The data generated by flex sensors in real-time is correlated with optimal draw length of the three bows determined during calibration. Based on data from flex sensors, accelerometer & anemometer, my app accurately calculated the desired draw length and angle to hit the target in a consistent way. The voice instructions were clear to aim the shots even when I blind-folded myself. My tool met the objective of incorporating my coach's guidance as scientifically calculated real-time feedback to effectively improve my game.

Conclusions/Discussion

I validated my project with Perkins & Texas schools for the blind, Professor Ting in the Vision Program at SFSU and with archery coaches for its effectiveness as a low-cost scientific tool to incorporate feedback during practice. In future I would like to incorporate target board scoring with image analysis so the game can be recorded, and blind people can know the score easily. As more users use my system, I would like to get more training data, to better my machine learning models for even better predictions in determining the accuracy of a shot.

Summary Statement

Created a low-cost archery assistant App for sighted & visually impaired archers that incorporates data from sensors to scientifically provide real-time, accurate feedback during practice for improved training.

Help Received

I designed and built the archery assistant myself. Learned Physics and Swift programming from my sister. Used BigML website to create models for accurate predictions and to evaluate results. Conducted all testing with arrows/target boards only in archery ranges under the supervision of my dad.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)

Iliana Close; Keira Swei

Project Number

J0803

Project Title

Automated Feature Detection for Diagnosing Neurodegenerative Diseases

Abstract

Objectives/Goals

The objective of this project was to determine if feature detection algorithms could identify neurodegenerative lesions in brain MRIs. Neurodegenerative lesions, including those found in patients diagnosed with Langerhans Cell Histiocytosis (LCH), are frequently missed through visual inspection of brain MRIs due to their low signal intensity. Furthermore, the detection process is difficult due to the high diversity of lesions and the similarity with normal tissue. By separating the brain into distinct regions, subtle changes in intensity between adjacent pixels in the image can be enhanced and used to identify lesions.

Methods/Materials

Brain MRI data were acquired from parents of patients that are members of a closed histiocytosis group on Facebook. A database of 107 reported cases of histiocytosis was compiled; of these, 27 separate brain MRIs were obtained from 13 patients. Two types of software were used to analyze the MRIs, including OsiriX for initial viewing and storage, and MATrix LABoratory (MATLAB) for feature/lesion identification and extraction. The algorithm that was developed was called Signal Characterization for Neurodegenerative Lesions (SCANL).

Results

Using SCANL, neurodegenerative lesions were detected in the cerebellum of 13 of the 14 MRIs previously identified by neuroradiologists, which is a 92.9% success rate. Furthermore, SCANL identified neurodegenerative lesions in an additional 3 patients (5 MRIs) from the Clean Scans group. Successful identification of lesions was independent of the magnetic field strength of the MRI machine and the size of the lesion, but dependent upon the intensity of the background cerebellum.

Conclusions/Discussion

SCANL was able to identify 92.9% of lesions previously found by neuroradiologists and lesions in an additional 5 MRIs from 3 patients who were previously told their scans were clean; 2 of these patients, after additional review by their oncologists, have now received independent verification of neurodegenerative changes. Based on these results, we recommend that doctors include SCANL (or feature detection algorithms) as part of the neuroradiology review of brain MRIs.

Summary Statement

We developed an automated feature detection algorithm to find neurodegenerative lesions in brain MRIs.

Help Received

My mother taught me how to code.

Name(s)
John Benedict A. Estrada

Project Number
J0804

Project Title
Comparison of Plant Chlorophyll Measurement Utilizing a Ground SPAD Meter vs. a Low Altitude Multispectral Camera

Abstract

Objectives/Goals
Measuring the chlorophyll content using a SPAD meter is tedious, laborious, expensive, and can potentially spread pests and diseases in the field. The aim of this study is to systematically obtain quantitative data to compare the efficacy of multispectral imaging techniques with ground-based SPAD meter readings and see if the vegetation indices obtained through small UAS photogrammetry correlate well with the standard ground-based chlorophyll measurement.

Methods/Materials
A field experiment on different N-fertilizer rates on Broccoli plants (Brassica oleracea cv. Marathon) was used in this study. A small unmanned aircraft system (UAS), commonly called a drone with a GPS-enabled multispectral camera was built. The UAS was flown autonomously at a low altitude (100 feet) and light reflectance values from the broccoli plants were recorded. These were used to calculate the vegetation indices: GNDVI, NDVI and CVI. The chlorophyll content of the broccoli plants on the ground were also measured using a standard SPAD 502 meter on the same day the images were taken. The vegetation indices were compared to SPAD measurements using regression analysis.

Results
The regression analysis showed that there are strong positive linear correlations between GNDVI, NDVI, and CVI with the SPAD readings. GNDVI, however, had the highest correlation (R2 = 0.92) compared to the other vegetation indices, NDVI and CVI, with R2 values of 0.74 and 0.7 respectively.

Conclusions/Discussion
This study showed that all three vegetation indices obtained using a UAS mounted camera and the technique developed in this study to exclude the soil pixels from plant pixels can reliably estimate the chlorophyll content in broccoli plants. It has the advantage of speed in obtaining measurements over large areas and avoids the potential spread of pests and diseases. By knowing the estimated chlorophyll content of the plants, farmers will know when and where it is necessary, and how much fertilizer to apply which can save money and protect our environment by avoiding ground water contamination.

Summary Statement
This project showed that a drone-mounted multispectral camera to calculate vegetation indices is as effective as a standard SPAD meter in measuring the chlorophyll content of broccoli plants.

Help Received
Dave Goorahoo, Ph.D., California State University - Fresno

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)

Lou Fowler

Project Number

J0805

Project Title

The Creation of a Hardy-Weinberg Calculator for Five Alleles Utilizing Python Coding Language

Abstract

Objectives/Goals

The objective of this project is to create a pair of calculators for the Hardy-Weinberg equations using Python coding language.

Methods/Materials

A computer and Python coding software were used in this project.

Results

This project was successfully completed. The calculators developed in this project enable the user to calculate the genotype and allele frequencies in a population with up to five alleles. These are the only Hardy-Weinberg equations calculators that can calculate from genotype to allele frequency.

Conclusions/Discussion

The goal of this project was to create a pair of calculators for the Hardy-Weinberg equations. The goal was achieved in this project. This project improves upon existing versions by allowing the user to calculate allele and genotype frequencies for a multi-allele gene. The Hardy-Weinberg equations are an inexpensive and noninvasive means of determining the frequency of a genetic disease in a given population. These calculators can also be used to easily calculate the frequency of other health-related genotypes. For example, the frequencies of blood types in a population can be found with these calculators. This is critical information for hospitals and blood banks. The function of these calculators is especially important with the growing understanding of the influence of genetics on health.

Summary Statement

I created a pair of calculators for the Hardy-Weinberg equations that can calculate both allele and genotype frequency in a population using up to five alleles.

Help Received

I did internet research on Python coding. I based my work on previous work done on programming quadratic equations. I was mentored by Tyler Sutterley, a software engineer. I wrote my own code and did all of my work at school or at home.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Saurav S. Gandhi

Project Number
J0806

Project Title
Eye Track: An Indoor Navigation System for the Visually Impaired

Abstract

Objectives/Goals
Most individuals can't detect physical obstacles and avoid them. For example, if a table is blocking one's path, most people would walk around it to get to the other side. But unfortunately, visually impaired people don't have this ability. This project focuses on creating a wearable, assistive technology, for the visually impaired, that allows them to navigate their indoor surroundings.

Methods/Materials
There have been many attempts to solve this problem with the use of sensors, vibrotactile and voice feedback, computer vision cameras, and RFID technology. The Eye-Track was created using a cheap computer vision camera (cmuCam5 Pixy Cam), vibration motors for vibrotactile feedback, and a fanny pack. This camera is capable of recording color signatures of objects and detects pre-programmed obstacles, by its hue. The Eye-Track was programmed to vibrate left or right, depending on the placement of the obstacle in the camera's field of vision, to aid the user in avoiding the obstacle. Eye-Track also incorporates an emergency button connected to the user's phone via Bluetooth, to contact a family member via SMS in case of an emergency.

Results
This product was tested in an obstacle course for the blindfolded user to walk through. The test criteria were based off on whether the user was able to identify the obstacle, avoid the obstacle, reach the destination, and make it through the obstacle course without touching a single obstacle. The success (%) was calculated for one, two and three obstacles. In summary, the "Eye-Track" product was 87.5% successful with one obstacle, 85% successful with two obstacles, and 72.5% successful with three obstacles.

Conclusions/Discussion
In conclusion, the success percentage decreased as the number of obstacles increased. A more sophisticated computer vision camera, with the ability to store more signatures and calculate depth would further this product. Additionally, there were numerous qualitative learnings related to the pace of walking, the distance between objects, and external light conditions. Improving these components could make the Eye-Track safer, and more apt for mass production.

Summary Statement
A wearable computer vision based indoor navigation system, for the visually impaired.

Help Received
I received help from my teacher, Ms. Scilingo who was my advisor for this project and guided my thinking as I developed the initial design of the wearable. She provided feedback on the multiple options I was considering. I also received help from my parents in funding the project, and providing support as I

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Madalynn L. Hein

Project Number
J0807

Project Title
Animated Storytelling with Alice 3 Programming

Abstract

Objectives/Goals
The purpose of this project is to learn to create animation using the Alice 3 programming language to create 3D figures that represents and demonstrates a story.

Methods/Materials
Computer and Alice 3 Programming IDE
Create some basic animation to learn how to use the Alice programming tool and language. Write a story I want to animate. Write code to animate the story and show to others to get feedback. Update as needed to make the story/animation better and repeat until finishes product matches the outlined story.

Results
Initial feedback was to have more of the story in the beginning to get background information earlier, the text displayed longer to make sure viewers have time to read all the text, something stated at the end as a "lesson learned" story, and different views of the character for the walking scenes. I updated the program based on this feedback and the result was a much better animated story.

Conclusions/Discussion
I learned how the Alice 3D animation programming language works. I also learned about the different animation orders and programming logic. The entire code by default goes in order so I had to add a "do at the same time" logic for certain parts such as walking. I learned a story can be animated but the length of the story needs to match the amount of time you have to create the animation for it. I had to shorten my story I wanted to animate in order to finish my project in time.

Summary Statement
I learned to use the Alice programming tool and logic loops to design and program a 3D animated story.

Help Received
I learned how to use the Alice programming tool on my own. I created the code myself but did need to look at examples posted by Anuka Dewahga for how to get characters to walk correctly. My mom also gave me feedback on the flow of the animation.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Peter M. Husman

Project Number
J0808

Project Title
A Computer Traffic Simulation: Vehicle Throughput across a Grid of Intersections Varying Driver and Car Parameters

Abstract

Objectives/Goals
This computer simulation was created to produce a more accurate model of traffic, taking into account realistic driver behaviors.

Methods/Materials
The simulation was written on a computer in C# and developed in Visual Studio 2017. The code was stored in a git repository and the changes were logged and managed with GitHub.
The simulation was analyzed by how many cars passed through the test environment as compared to the number of car crashes across hundreds of trials with varied car and environmental parameters. The cars' parameters included an acceleration and deceleration constant, a maximum speed, and a target headway. Each intersection is assigned light timing at the beginning of each trial.

Results
The trials were able to approach high car survival rates, in some cases almost 90%, as the algorithm was improved over time. The throughput percentage for the simulation varied slightly in each trial, as the parameters of each car and intersection were randomized to help increase realism.

Conclusions/Discussion
The simulated cars were able to replicate some behaviors seen in real cars with real drivers. According to the research, drivers on a road without a bottleneck exhibit three behaviors: a period of acceleration, a period of stability, and a period in which the cars slow down in a jam. In the simulation, these behaviors also exist in a very visible form.

Summary Statement
This computer simulation of traffic modeled cars as independent units, each implementing a basic algorithm based on internal and environmental factors, to achieve somewhat realistic behaviors to eventually predict routes and flow.

Help Received
I wrote and executed the simulation and the data analyzer by myself. Stan Khaykin taught me the concept of linked lists.

Name(s)
Nithika Karthikeyan

Project Number
J0809

Project Title
The Design of Algorithms to Encode English Text in Amino Acids Using Digital Data Compression Techniques

Abstract

Objectives/Goals
Worldwide digital data is forecast to grow to 160 zettabytes (10^18 KB) in 2025. Traditional storage solutions are not keeping up with this exponential demand, increasing the cost of data storage. Most data (75%) is archival in nature, for which storage access is not time critical.
DNA and amino acids are being considered as unconventional, high-density storage media for archival data. Their small size (sub-nanometer for DNA and nanometer for amino acids) and non-binary nature can result in 85% space savings compared to current storage solutions.
The goal of this project is to devise algorithms to further increase the space efficiency of amino acid storage media using data compression techniques and data characteristics.

Methods/Materials
Two encoding and decoding algorithms were invented. The FixedLength algorithm assigns two amino acids to a byte of data, based on size. The VariableLength algorithm uses Huffman encoding and text characteristics, to assign one or two amino acids for every byte.
Using Pascal and a Windows PC, the algorithms were implemented. The output of encoding, a sequence of amino acids (peptide), was checked for stability and structure using free simulation software. Peptide sequencing to ensure it stored the sequence correctly, was done using free Mass Spectroscopy simulation. These algorithms were compared to the baseline algorithm, Sabry. Each algorithm was run over nine English texts of differing lengths. The decoded text was compared to the original for correctness. An end-end amino acid storage system has been proposed, but not implemented, as it is outside the scope of this project.

Results
FixedLength is 35% better than baseline and VariableLength is 42% better. The encoding and decoding times of both algorithms are comparable to Sabry. Overall, both algorithms proved to be more efficient than baseline.

Conclusions/Discussion
Computations showed that VariableLength can store "The Bible" in 8% of the space needed by solid state memory. This will bring huge space savings at data centers and the electricity needed to run them, thus reducing storage cost. As slow access of archival data is acceptable, chemical processes involved would not deter the application of this storage solution.
Enhancements could be made to represent other media types. Duplicate amino acids when sequencing has to be resolved. Actual synthesis and sequencing would prove the system works.

Summary Statement
This project involves the creation of two algorithms that reduce data storage costs using unconventional storage media.

Help Received
I designed the algorithms on my own, but received some help with the understanding of amino acids from BioCurious as well as several online websites.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Rohan A. Karunaratne

Project Number
J0810

Project Title
A Real-Time Tactile Image Generator for the Blind and Visually Impaired

Abstract

Objectives/Goals
The goal of this project was to successfully create tactile image by taking a digital image, breaking it up into a 6 by 6 array, and then assigning each array to one push rod which will go up and down based on how far away the object is. There by allowing a blind person to "feel" an image of his surroundings in real-time. The purpose of this project was to help blind and visual impaired people "see" their surroundings by creating a tactile representation of their surroundings. With the help of this technology, the blind and visually impaired navigate more efficiently through the world.

Methods/Materials
Materials:
SG90 Hobby Servos (x36), 3D printer, Computer with Python and Arduino programming software, Spring Pins, Arduino Uno, PCA9685 servo driver, Jumper Wires, Camera.
Procedure:
1. Design the fixture to hold the servos, the push rods, and pistons
2. 3D print the designs using the 3D printer
3. Install the servos and push rods so the fixture works
4. Take an image from a camera. Write python code to break the image up into a 6 by 6 array. Depending on the luminosity value of each sector, make the pistons go up and down using Arduino code. The Arduino should communicate with the PCA via I2C to drive the servos.
5. After working with a still image, write python code to take live video to pixelate the video into a 6 by 6 array that can continuously update depending on the video.

Results
I plotted a chart that shows the extension of the push rod vs the luminosity values. In the chart, a blue line represents the theoretical position the push rod should take, while an orange line shows were the measured extension. The theoretical equation is as follows: 10mm*(1-COS(3.14*Luminosity Value)/255)).

Conclusions/Discussion
Overall, this project can help the blind navigate through their every day lives. If I can reduce the size of my project, blind people can use this in their everyday lives.

Follow-up Project:
While light and dark represent a visual image, the actual contours of the object may make more sense to a blind person feeling the object because light and dark are not a full representation. It may be better to

Summary Statement
In this project, I created a machine that could take images from a camera and make them into a tactile image with applications for the blind and visually impaired.

Help Received
I got help from my cousin with some of the code regarding com ports. I used my dad's 3d printer to make my parts. The assembly, 3d designs, and majority of the code was done by me.

Name(s)
Houjun (Jack) Liu

Project Number
J0812

Project Title
Comparing the Effects of Various Corpora's Qualities on NLG/NLP Systems

Abstract

Objectives/Goals
The experiment is designed to analyze the relationship between the input quality of a dataset and the output quality of a dataset after it is processed a Natural Language Generation (NLG) system - a system that would generate text based on an input. During the experiment, the focus is to identify whether or not the input/output qualities affect each other in a linear pattern (hence, having an equal ratio in growth.)

Methods/Materials
Four corpora of incrementing quality, two generation systems, Python and Wolfram Mathematica -- where scripts written for this experiment computation is executed, Gold Standard benchmark corpus "1200 Graded Sentence for Analysis" for evaluation.

The four corpora are all separately fed in small batches into the two generation systems, which generates five batches of 100 sentences based on the corpora. These sentences are evaluated according to the benchmarks of the Gold Standard corpus. An average quality score between 1-5 is found for each batch. A best fit line between the input quality of each corpus against the average quality score across all five batches for that corpus is found, and finally, the error rate of that fit line is identified to test the linearity of the dataset.

Results
According to the data, both model's best fit lines all have near-zero error values. Furthermore, the ratios between input and output qualities on each of these models stayed approximately the same around 1.

Conclusions/Discussion
Since both sets of numerical data collected above identified a linear pattern between the input quality and the output quality of the NLG system, it is shown that the data observed had a linear relationship with each other. This result would help model the performance of an NLG system for the future.

Summary Statement
NLG systems' input and output corpora's qualities are compared, and it is found that, when data of various quality is fed through and NLG system, the input and output data had a linear relationship in quality.

Help Received
My science teachers helped proofread the display board and check for semantic and grammatical errors.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Kavya M. Pandrangi

Project Number
J0813

Project Title
Automated Insulin Delivery Model with Feedback for Diabetics: A Microcontroller and Sensor Based Bioengineering Project

Abstract

Objectives/Goals
Many diabetics constantly monitor their blood sugar levels and administer insulin as needed. To help diabetics, extensive research is being done to create a device which functions like pancreas by automatically administering insulin as needed, but many obstacles stand in the way of adoption - e.g., trust issues. The goal of my project is to design a contraption to simulate artificial pancreas which will also provide real-time feedback to the users to alleviate trust issues.

Methods/Materials
I used acid-base neutralization concept to simulate the process of insulin normalizing high blood sugar levels using vinegar in place of blood and baking soda for insulin. First part of my design consists of a DC motor configured to pump baking soda solution when the vinegar solution is acidic enough. A conductivity sensor is placed in the vinegar jar to work like a variable resistor as the baking soda is pumped. The potentiometers in my circuit are calibrated such that the pump would stop pumping baking soda as soon as the vinegar is neutralized. To provide feedback to the user, I created a mobile application that takes input from an arduino connected to an ultrasonic level sensor. I connected ultrasonic sensor to the top of the baking soda solution jar so I can measure the amount of baking soda left in the jar as it is being pumped into the vinegar jar. The Arduino then publishes the information to the mobile app about when the baking soda was pumped and how much is left. When the level of the baking soda is low enough, the arduino will send out a notification to the app. This would alleviate the user engagement and trust concerns since the user would be getting constant feedback whenever insulin is pumped.

Results
After I calibrated the potentiometers to stop the pump when the vinegar solution is neutralized, I was able to accurately determine how much baking soda solution (i.e., insulin in the model) was used. I was then able to publish this information to the mobile app as well as to determine when the insulin reservoir needs to be refilled or replaced.

Conclusions/Discussion
This model shows that automation and real-time monitoring of insulin delivery are feasible. Some enhancements are still necessary which I plan to address in future: miniaturization of the circuits, more accurate ways to determine insulin levels in the pump (instead of ultrasonic sensor) and publishing blood sugar levels also to the application.

Summary Statement
My project is about building an automated insulin delivery model with real-time mobile app based feedback for diabetics using a microcontroller and ultrasonic and conductivity sensors.

Help Received
I learnt about acid-based neutralization and electronic circuit concepts at school. I then designed and programmed the model after reading thru online material on circuits, arduino and sensors. My mom reviewed my circuit and dad reviewed my code.

Name(s)
Sneha Revanur

Project Number
J0814

Project Title
Analyzing Gender-Based Violence and Aggressive Behavior through Social Media Data

Abstract

Objectives/Goals
My project goal was to develop a computational model to classify Tweets for relevance to gender-based violence (GBV), a prevalent human rights issue that transcends a variety of demographics.

Methods/Materials
I defined four classes that pertain to GBV: Physical Violence, Sexual Violence, Harmful Practices, and Other. I used the Python programming language. I extracted data from the Twitter API based on class-specific search criteria and employed the Natural Language Toolkit (NLTK) for natural language processing on Tweet text. Of the 4,000 filtered Tweets, 80% were used as training data and the rest were used as testing data. I used the Naive Bayes classification algorithm to train the machine learning model. I went on to conduct a comparative analysis of two feature sets consisting of unigrams and bigrams. I also constructed a confusion matrix to better analyze the model's performance.

Results
The feature set based on my search criteria had the highest accuracy, with over 85%. For the NLP-based features, Harmful Practices had the highest precision and Other had the lowest. For the search criteria-based features, Harmful Practices had the highest precision and Physical Violence had the lowest. The countries that most frequently discussed GBV included the US, UK, Canada, India, and Australia.

Conclusions/Discussion
I was able to meet my project goal, and successfully leveraged computational linguistics, machine learning, and computational social science to develop a highly accurate Tweet classification model for GBV.

Summary Statement
I developed a computational model that can independently classify Tweets into one of four classes germane to gender-based violence (GBV) while harnessing natural language processing and machine learning.

Help Received
I developed this model myself with the support of my parents and teacher sponsor.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Holden T. Robertson

Project Number
J0815

Project Title
Which Cipher Is the Most Difficult for a Computer to Decode?

Abstract

Objectives/Goals
The objective of my project was to determine which type of cipher was the hardest for a computer to decode: The Ceaser Cipher, The Polybius Square, and The Vigenere Cipher.

Methods/Materials
To execute my project I created different files to hold each of my cipher codes. Inside each file, I first started a timer that I had imported from a dictionary. I wrote code to loop through each character in the ciphertext. Each cipher required different code to decrypt the characters. For the Vigenere Cipher, I had the code build a matrix for me. For the Polybius Square, I manually created a matrix. For my Caesar Cipher, I created a dictionary to hold all of my letters and their values. Also, I had to loop through all the possible different shift values. Lastly, I stopped my stopwatch and recorded my elapsed time in nanoseconds. Then I divided my elapsed time by 1,000,000,000 to get seconds. The materials that I used are the following: a stopwatch that was built into my code, and by visually assessing the number of lines of code.

Results
The results showed that the Vigenere Cipher had the greatest lines of code, with the Polybius Square coming in second, and the Caesar Cipher in third. For the overall time the Caesar Cipher comes in first with the longest time, Vigenere Cipher in second with the next longest, and Polybius Square with the least.

Conclusions/Discussion
I believe that I received these results because the Vigenere Cipher had the most complex code, due to its 26 by 26 matrix, so it would have the most lines of code. The Polybius Square has the second greatest lines of code because it has a matrix that is much smaller than the matrix in the Vigenere Cipher but still more complex than the Caesar Cipher. The Caesar Cipher has the least lines of code but the most time because it is repeating a fairly simple task many times. The Polybius Square received the least time because it had the least number of steps of all the codes. The Vigenere Cipher had the second least time because it was doing a fairly simple task that had complex components.

Summary Statement
I showed that the Vigenere Cipher was the most difficult due to it's 26 by 26 matrix, which took the most lines of code.

Help Received
I would also like to thank my father for instructing me during the allotted amount of time while I am creating the computer code to decode the cipher.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Dana S. Saeteurn

Project Number
J0816

Project Title
How Sound Cues Affect a Visually-Impaired Person's Ability to Play Video Games

Abstract

Objectives/Goals
The objective of my project is to design a video game with sound cues so that a visually impaired person can have the ability to play it.

Methods/Materials
Developed "Help Cat Find His Clarinet" game on https://scratch.mit.edu/projects/143658317/ by incorporating clarinet music notes to guide players in achieving the game's objective, tested visually impaired players and players with no visual impairments of the same ages by timing how long it took to complete the objective of the game with and without blindfolds and with and without sound, interviewed/surveyed participants' experiences playing the game, and calculated score averages and analyzed data to see if hypotheses were correct.

Results
The test subjects strongly agreed that the sound cues were helpful when you are blind because Statement #3 (The sound cues were helpful in the game.) on my survey sheet got the highest average rating of 4.6 out of 5.0. Also, the survey revealed that the test subjects strongly disagreed with Statement #2 (It was possible to meet the objective of the game with no sound cues.) with a rating of 2.6 out of 5.0. In addition, only 3/10 kids were able to complete my game without sound cues within a 1:30 min. time limit. Lastly, the visually impaired kids finished my game faster at an average speed of 55 seconds while it took the seeing kids an average of 59.07 seconds to finish.

Conclusions/Discussion
My experimental results support my hypotheses that sound cues are helpful to the visually impaired and that the visually impaired test subjects can complete the video game faster than the test subjects can with no impairments. This shows that visually impaired kids rely heavily on hearing to learn. Since a visually impaired person can play video games like a regular person can, then we should understand that all people will benefit from having sound to assist them in their daily activities. Most importantly, having a disability shouldn't prevent someone from doing what an able body can do.

Summary Statement
My project is a video game enhanced by sound cues that I coded to allow the visually impaired to be able to play it.

Help Received
I designed and created the video game by myself, but I learned the basics of coding at the UC Merced Mother and Daughter Science Camp. The Merced County Office of Education helped me get in contact with visually impaired students throughout the county so I can be able to perform my experiment.

CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)
Ziyad Soliman

Project Number
J0817

Project Title
Map It Right!

Abstract

Objectives/Goals
The motive of this project is to create an autonomous robot using the LIDARLite V3 sensor to produce a 2D map of an uncharted area. A secondary application enables the robot to measure distances.

Methods/Materials
The robot was constructed from a LIDARLite V3 Sensor, AX-12A Servo, Arduino Board, Breadboard, and two KR-250 Servos. Downloaded a software called Arduino to program. The robot was programmed to map out an area and then compared to the actual room.

Results
The first several trials did not generate an accurate map as the corners were sharp and the lines were crooked. Changing the different connections on the robot and modifying the code produced accurate results for the ten rooms tested.

Conclusions/Discussion
The implications of my results may be used in numerous situations, specifically the military, which helps soldiers have a plan before going in an unknown area. I created my robot to have the ability use various sensors for many other missions, such as locating radiation hotspots in a nuclear power plant or locating rust on a bridge.

Summary Statement
I created an autonomous robot to map out an unknown room or area using a LIDARLite V3 sensor.

Help Received
My older brother recommended I learn the Extended Kalmin Filter (EKF) to help with my calculations and my teacher reviewed my results.

Name(s)
Joseph P. Thomas

Project Number
J0818

Project Title
A Machine Learning Approach to Predicting NBA Rookie Potential

Abstract

Objectives/Goals
The objective of this project was to apply machine learning algorithms to predict whether an NBA rookie will become an all-star based on his first year statistics.

Methods/Materials
The data obtained from Kaggle, a data science competition platform, contained per-year statistics for every NBA player from 1950 to 2017. Another dataset containing every player who became an all-star was obtained from Wikipedia. These two datasets were combined into one with 52 columns and 3,921 rows. Python was used to trim the data by removing the years 1950 to 1982 because most statistics were not recorded during these years. The years 2006 to 2017 were also removed to ensure that every rookie had finished his career. Seventy percent of this remaining dataset was devoted to training and the rest was used for testing. The training set had the problem of class imbalance, which in this case means that there are many more non all-stars than all-stars. This problem was solved by removing random non all-star players so that the number of non all-stars was equal to the number of all-stars. The rookie dataset was then classified by four machine learning algorithms that were already trained.

Results
Each machine learning algorithm was run three times on the test set and an average was calculated for the accuracy, precision, recall, and F1 score. The Naive Bayes classifier performed the best with an average accuracy of 82%, precision of 33%, recall of 80%, and a F1 score of 47. These results established viability of machine learning as a way to overcome cognitive biases in player-trading decisions.

Conclusions/Discussion
The Gaussian Naive Bayes classifier performed much better than other algorithms because it works very well with small datasets and makes probability based decisions. The Naive Bayes classifier was correct in its predictions 33% percent of the time, which is better than a classifier that always assumes one answer. Removing the class imbalance contributed to the success of this project by avoiding the fixed answer, non all-star.

Summary Statement
Using analyzed data and the power of machine learning, I discovered that computer algorithms can assist general managers to overcome cognitive biases when making player-trading decisions.

Help Received
I applied machine learning algorithms to the data that I gathered and analyzed. The help that I received was from my science teacher, Mr. Colucci to understand the capabilities of the different machine learning algorithms.