VAAL UNIVERSITY OF TECHNOLOGY

Inspiring thought. Shaping talent.

AIBY3A/AIBU3A

BUSINESS ANALYSIS 3.2 AI PROJECT

AI ANIMALS MONITOR

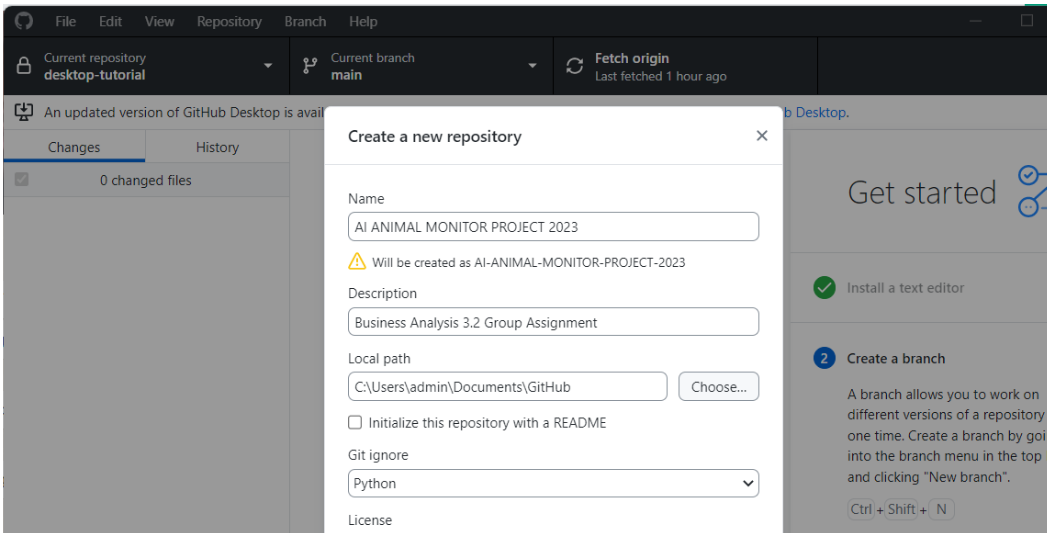
GROUP MEMBERS:

Student ID:	Surname & Initial:	Signature:
220110492	Davhana T	T.Davhana
221194126	Tsatsane OT	OT.Tsatsane
219155534	Mabena MD	MD.Mabena
219147191	Maisela N	N.Maisela
221232249	Xekwa T	T.Xekwa
220173435	Mabunda B	B.Mabunda

Lecturer: Mr Polyose X and Mr Matsela

Submission Date: 06 Sep 2023

Contents



File Edit View Repository Branch Help
Current repository
desktop-tutorial
Current branch
main
Fetch origin
Last fetched 1 hour ago
An updated version of GitHub Desktop is avail
b Desktop.
Changes
History
0 changed files
Create a new repository
Name
AI ANIMAL MONITOR PROJECT 2023
Will be created as AI-ANIMAL-MONITOR-PROJECT-2023
Description
Business Analysis 3.2 Group Assignment
Local path
C:\Users\admin\Documents\GitHub
Choose...
Initialize this repository with a README
Git ignore
Python
License
Get started
Install a text editor
Create a branch
A branch allows you to work on different versions of a repository one time. Create a branch by goi into the branch menu in the top and clicking "New branch".
Ctrl + Shift + N

AI Solution

The technologies of the Fourth Industrial Revolution (4IR), including Artificial Intelligence, Augmented Reality, Robotics, and 3-D printing, are rapidly transforming the manner in which individuals generate, share, and distribute value. In response to this transformative shift, we have taken the initiative to develop an Artificial Intelligence system known as Smart-Shepherd for the Vaal Community. It's important to note that 4IR is not merely a forecast of the future; it represents a compelling call to action. It presents a vision for the development and governance of technologies in a way that promotes a more empowering, collaborative, and sustainable framework for social and economic progress. The Smart-Shepherd system has been designed to assist in the detection and monitoring of livestock, ensuring their well-being and safety.

The Machine will perform the following tasks:

Identifying various objects, such as differentiating between types of animals and humans, while also ensuring that animals pose no threat to each other is a primary function of this robot. It's important to emphasize that the robot's intention is not to supplant Human-Shepherds but rather to assist and collaborate with them in the responsible management and well-being of livestock.

Problem definition

Rhinos and other animals are being killed at an alarming rate, and despite the statistics, there has been no solution to this pressing issue. After conducting research in our community, we found that shepherds are just as afraid of omnivores as they are of carnivores. Therefore, it's crucial to have a robot that can take care of livestock.

While human shepherds are doing a great job, they sometimes fall sick and are unable to work in extreme weather conditions. A shepherd robot would be very useful on farms as it can easily interact with livestock and understand how animals communicate with each other. Additionally, this robot can regulate water consumption, ensuring that flocks do not consume excessive amounts of water, since it will be easily maintained.

Benefits of our solution

The following are some potential uses of technology in the agriculture industry:

- Detecting diseases in animals
- Counting livestock
- Identifying misplaced livestock
- Reducing livestock poaching
- Ensuring the safety of human workers
- Detecting weather patterns
- Creating job opportunities for IT specialists
- Reducing the risk of sickness and fatigue among workers

Type of intelligence

Our robot comprises of Linguistic and spatial intelligence, which involves the capacity to use language to accomplish the goal of detecting livestock's.

Machine learning:

It is worth noting that the semi-supervised learning approach can be highly effective for robots, as it provides the ability to compare datasets, reinforce mutual observations, and correct any errors or over-generation. Ultimately, this can help the humanoid to adapt and balance strategies in a more efficient manner.

Learning Approaches:

Below is a list of some important concepts that we will be using in our system:

- **Kinematics**: This will help our system understand the motion of objects in the body.

- **Bayesian Models**: This involves formulating subjective probabilities to express existing information, carefully modelling data structure, and allowing a utility function to express how the value of each alternative decision is affected.

- **Support Vector Machines**: This is a type of supervised machine learning that will help us analyze data for classification and regression analysis.

- **Unsupervised K-Means Clusterin**g: We will be using this to group animals into distinct categories based on their similarities and differences.

Data:

Data Wrangling:

Firstly, we are going to try and explore the data to check for missing values/erroneous

Entries and comment on redundant features and add additional ones, if needed.

- Data types
- No null values
- Imbalanced classes of predicted variable
- Transform variables into binary

Exploring the Data:

- Explore the distinctive features of the data
- Determine how good a feature it is for prediction

Baseline Modelling:

We are going to start modelling to learn more above our variables! For this first run we are going to use ALL our non-categorical variables.

We used following list of classifiers used to predict the model accuracy:

- Logistic Regression
- Random Forest
- Support Vector Classifier
- Class Misbalancing
- Decision Tree

A1 | ANIMAL_NAME

	A	B	C	D	E	F	G
1	ANIMAL_NAME	ANIMAL_ID	ANIMAL_TYPE	IMAGE	ANIMAL SIZE	COLOR	AGE
2	LION	1001	CARNIVORE	LION.JPG	190 KG	GOLD	7
3	ZEBRA	1002	HERBIVORE	ZEBRA.JPG	300KG	BLACK & WHITE	7
4	ELEPHANT	1003	HERBIVORE	ELEPHANT.JPG	6 000 KG	GREY	10
5	RHINOCEROS	1004	HERBIVORE	RHINO.JPG	1 000 KG	GREY-BLACK	7
6	LEOPARD	1005	CARNIVORE	LEOPARD.JPG	90 KG	BROWN	6
7	GIRAFFE	1006	HERBIVORE	GIRAFF.JPG	500KG	LIGHT BROWN	5
8	HIPPO	1007	HERBIVORE	HIPPO.JPG	1200 KG	BLACK	4
9	BUFFALO	1008	HERBIVORE	BUFFALO.JPG	800KG	BLACK	10
10	OSTRICH	1009	HERBIVORE	OSTRICH.JPG	100 KG	BLACK & WHITE	4
11	IMPALA	1010	HERBIVORE	IMPALA.JPG	70KG	BROWN	8
12	SPRINBOK	1011	HERBIVORE	SPRINBOK.JPG	60 KG	BROWN	6
13	CHEETAH	1012	CARNIVORE	CHEETAH.JPG	64KG	TAN WITH BLACK SPOTS	7
14	JACKAL	1013	CARNIVORE	JACKAL.JPG	10KG	BROWN	8
15	WARTHOG	1014	OMNIVORE	WARTHOG.JPG	60KG	DARK RED-BROWN	5
16	HEDGEHOG	1015	HERBIVORE	HEDGEHOG.JPG	500 G	BROWN	7
17	CAPE GRASBOK	1016	HERBIVORE	GRASBOK.JPG	120 KG	BROWN	6
18	ELAND	1017	HERBIVORE	ELAND.JPG	600 KG	LIGHT BROWN	8
19	NYALA	1018	HERBIVORE	NYALA.JPG	88KG	DARK BROWN	9
20	ANTELOPE	1019	HERBIVORE	ANTELOPE	900 KG	BROWN GOLDEN	9
21	TYGER	1020	CARNIVORE	TYGER.JPG	180 KG	ORANGE-ISH	8
22	DEER	1021	HERBIVORE	DEER.JPG	50 KG	BROWN	5
23	CHIMPANZEE	1022	OMNIVORE	CHIMPANZEE.JPG	35 KG	BLACK	3
24	CROCODILE	1023	CARNIVORE	CROCODILE.JPG	900 KG	GREY	4
25	WILDEBEEST	1001	HERBIVORE	WILEBEEST.JPG	190KG	BROWN	6
26	GORILLA	1001	OMNIVORE	GORILLA .JPG	90KG	BLACK	6
27	BABOON	1001	OMNIVORE	BABOON.JPG	25KG	BLACK	4
28	GREATER KUDU	1001	HERBIVORE	GREATERKUDU.JPG	230KG	BROWN	7
29		1001		.JPG			
30		1001		.JPG			

Natural Language Processing:

NLP will be used so that it can engage carefully with the humans. We will teach it to understand the language we use to communicate by training it. The collected words, phrase or sentences will be analysed using lexical analysis. Reason for choosing lexical analysis that It will help identify and analyse the structure of words, there after there will be a division of chunk of text into paragraphs, words, phrase and sentences, then they will be documented by matching similar document of texts, content indexing and content summarization

The Google voice recognition API will be implemented for this model. Speech recognition makes communication possible between humans and computers by detecting changes in the sequence of words that make up a written text, which may then be processed and responded to spoken commands. Google speech recognition already has in-built steps used to getting text from users, decoding it, structuring it, and transforming it so that it can be understood by the machine, therefore, we do not need to worry about our model not recognizing short phrases, vocabularies, and long phrases. Limited factors would be the background sounds, but they can be avoided by limiting background noise when we interact with the machine.

Deep Learning:

Deep Learning is a branch of Machine Learning that allows machines to learn and adapt without relying on explicit instructions. This is achieved through the use of algorithms and statistical models, which analyze and draw inferences from patterns in data.

Our humanoid will use this approach to generate its own training data and improve its performance. It will capture data from close range to interpret long range sensor data and analyze rough terrain in 3D-scene analysis to identify mathematical subjects and physical science.

The machine will use sensors, cameras, and laser points to detect surroundings. A Deep Neural Network will create high-quality artistic images and separate/recombine image styles using a neural algorithm. This is because DNN is most powerful in images, and deals with pictures and faces on a daily basis.

To train the machine for object recognition, the DNN will use Convolutional Neural Networks and develop a representation of the input image. The input image is transformed into actual content compared to its detailed pixel values.

We can visualize the information from each layer of the network by reconstructing the image only from the feature maps in that layer. Higher layers in the network will capture high-level content in terms of objects and their arrangement in the input image.To obtain a representation of the style of input image, we use a feature designed to capture texture information, this feature space is built on top of this filter responses in each layer of the network. It consists of the correlations between different filter responses over the spatial extent of the feature maps by including the filter of multiple layers.

Finally, we obtain multi-scale representation of the input image, which captures its texture information but not global arrangement.

Libraries

In [453…

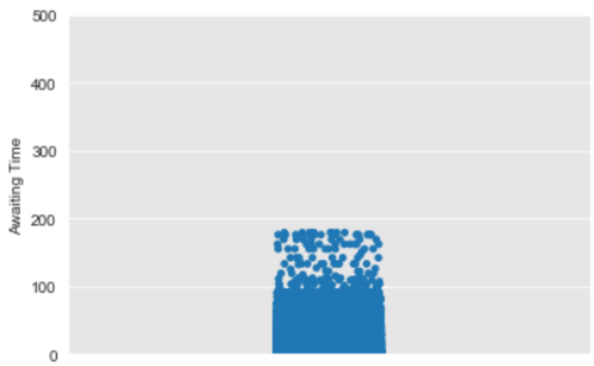
```
import pandas as pd
import numpy as np
import itertools
import matplotlib.pyplot as plt
import seaborn as sns
import datetime
%matplotlib inline


from sklearn.model_selection import train_test_split
from sklearn import ensemble
from sklearn.metrics import classification_report
from sklearn.model_selection import cross_val_score
from sklearn.metrics import confusion_matrix
from sklearn.svm import SVC
from sklearn.tree import DecisionTreeClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import roc_curve, auc
from sklearn.metrics import confusion_matrix
```

Out[544…

```
array(['F', 'M'], dtype=object)
```

In [545…

```
# Checking for outliers in AwaitingTime
sns.stripplot(data = df, y = AwaitingTime, jitter = True)
plt.ylabel('Awaiting Time')
plt.ylim(0, 500)
plt.show()
```



In [546…

```
df.ScheduledDay = df.ScheduledDay.apply(np.datetime64)
df['NoShow'] = df['No-show']
def calculateHour(timestamp):
    timestamp = str(timestamp)
    hour = int(timestamp[11:13])
    minute = int(timestamp[14:16])
    second = int(timestamp[17:])
    return round(hour + minute/60 + second/3600)

df['HourOfTheDay'] = df.ScheduledDay.apply(calculateHour)
```

In [547…

```
def probStatus(dataset, group_by):
    df = pd.crosstab(index = dataset[group_by], columns = dataset.NoShow).reset_index()
    df['probShowUp'] = df['No'] / (df['No'] + df['Yes'])
    return df[[group_by, 'probShowUp']]
```

Logistic Regression

Fit a model

In [484…

```
logreg = LogisticRegression(fit_intercept = False, C = 1e12, solver = 'lbfgs')
model_log = logreg.fit(X_train, y_train)
print(model_log)
print('Training set score:', logreg.score(X_train,y_train))
```

```
LogisticRegression(C=1000000000000.0, class_weight=None, dual=False,
                   fit_intercept=False, intercept_scaling=1, l1_ratio=None,
                   max_iter=100, multi_class='warn', n_jobs=None, penalty='l2',
                   random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
                   warm_start=False)
Training set score: 0.7904909467208017
```

Predict

In [485…

```
y_hat_test = logreg.predict(X_test)
y_hat_train = logreg.predict(X_train)
```

How many times was the classifier correct for the training set?

In [486…

```
#Initial Evaluation

residuals = y_train - y_hat_train
print(pd.Series(residuals).value_counts())
print(pd.Series(residuals).value_counts(normalize=True))
```

```
 0    69896
 1    17503
-1     1022
Name: No-show, dtype: int64
 0    0.790491
 1    0.197951
-1    0.011558
Name: No-show, dtype: float64
```

Support Vector Classifier

In [420…

```
from sklearn.svm import SVC
svc = SVC(kernel='linear')
svc_model = svc.fit(X_train,y_train)
print(svc_model)
print('Training set score:', svc_model.score(X_train,y_train))
```

```
SVC(C=1.0, cache_size=200, class_weight=None, coef0=0.0,
    decision_function_shape='ovr', degree=3, gamma='auto_deprecated',
    kernel='linear', max_iter=-1, probability=False, random_state=None,
    shrinking=True, tol=0.001, verbose=False)
Training set score: 0.797129641148596
```

In [422…

```
SVC_score = cross_val_score(svc, X_test, y_test, cv=5)
print('\nEach Cross Validated Accuracy score: \n', SVC_score)
print("\nOverall Logistic Regression/Support Vector Classifier Accuracy: %0.2f (+/- %0.2f)\n" % (SVC_score.mean(), SVC_score
```

```
Each Cross Validated Accuracy score:
 [0.80167345 0.80185478 0.80185478 0.80185478 0.80185478]

Overall Logistic Regression/Support Vector Classifier Accuracy: 0.80 (+/- 0.00)
```

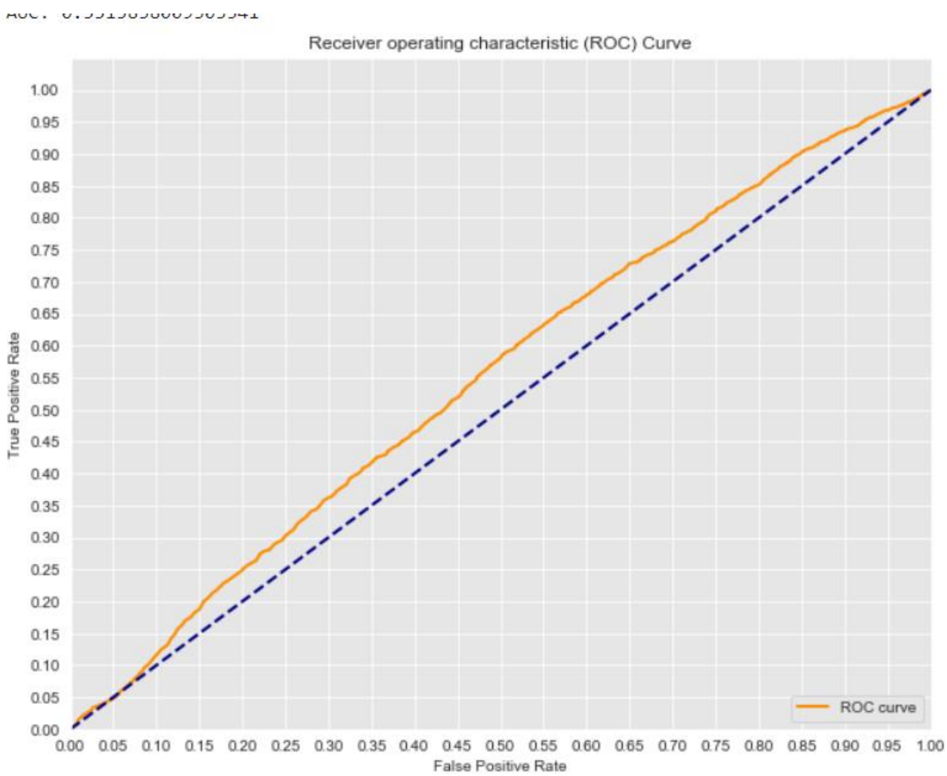
Create an initial model

In [490…

```
#Initial Model
logreg = LogisticRegression(fit_intercept = False, solver= 'lbfgs')

#Probability scores for test set
y_score = logreg.fit(X_train, y_train).decision_function(X_test)
#False positive Rate and true positive rate
fpr, tpr, thresholds = roc_curve(y_test, y_score)

#Seaborns Beautiful Styling
sns.set_style("darkgrid", {"axes.facecolor": ".9"})

print('AUC: {}'.format(auc(fpr, tpr)))
plt.figure(figsize=(10,8))
lw = 2
plt.plot(fpr, tpr, color='darkorange',
         lw=lw, label='ROC curve')
plt.plot([0, 1], [0, 1], color='navy', lw=lw, linestyle='--')
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.05])
plt.yticks([i/20.0 for i in range(21)])
plt.xticks([i/20.0 for i in range(21)])
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver operating characteristic (ROC) Curve')
plt.legend(loc="lower right")
plt.show()
```

AUC: 0.5515838009303541

Receiver operating characteristic (ROC) Curve



References

1. Jeju Media. (Year not provided). "5 Benefits of Bringing Robots in the Agricultural Sector." [Online] Available at: https://www.jejumedia.com/5-benefits-of-bringing-robots-in-the-agricultural-sector/ (Accessed: September 5, 2023).
2. Britannica. (Year not provided). "The Fourth Industrial Revolution." [Online] Available at: http://www.britannica.com/topic/The-fourth-Industrial-Revolution2119734 (Accessed: September 12, 2023).
3. Unknown Author. (Year not provided). "How to Build a Neural Network for Voice Classification." [Online] Available at: https://towardsdatascience.com/how-to-build-a-neural-network-for-voiceclassification-5e2810fe1efa (Accessed: August 28, 2023).
4. Grinnell, R. &. & Unrau, Y., 2008. *Social Work Research and Evaluation: Foundations of Evidence-based Practice.* New York: Oxford University Press.
5. Ha, G. & Ngo, T.C.T., 2021. Challenges in learning listening comprehension via Microsoft Teams among English majors at Van Lang University. *International Journal of TESOL & Education,* Volume 1(3), pp. 42-175.
6. Harris, J., Jones, M.L., Bailey, G.K. & Westbrook, R.F., 2000. Contextual control over conditioned responding in an extinction paradigm. *Journal of Experimental Psychology: Animal Behavior Processes,* Volume 26(2), p. 74.
7. Johannes , K., Daniela , . J. J.-B. & Glutsch, N., 2020. Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *European Journal of Teacher Education 43:4,* Volume 43:4, pp. 608-622.
8. Keerio, M.U., et al., 2022. Evaluating Students’ Perceptions of Microsoft Teams for Online Academics Improvement. *Pakistan Journal of Engineering and Technology,* Volume 5(1), pp. 56-67