GC-303

DEMPSTER-SHAFER THEORY OF COMPUTATION

Abstract

Python, a programming language, is used throughout this report paper in order to explain the Dempster-Shafer Theory. This report includes a literature review, an implementation process, a methodology, analysis data, and suggestions. The purpose of this report is to understand and make a clear knowledge of the issue that was addressed. Another objective of this paper is to comprehend the entirety of the Dempster-Shafer Theory and to implement it utilizing the Python programming language.

Introduction

The DST is a mathematical theory regarding proof that was devised by Stanley Dempster and Edward Schaffer, who are both psychologists. The Dempster-Shafer theory is a generalization of probability theory. It differs from traditional probability theory in that possibilities are often assigned to sets rather than to singletons that are mutually exclusive. Furthermore, it can be described in terms of a finite discrete space. In classical probability theory, evidence is never connected to more than one occurrence at a time (Xiao, 2020). The DST makes it possible to connect pieces of data to a variety of possible outcomes, sometimes known as "sets of occurrences." Therefore, without making any assumptions about the happenings contained within the evidentiary set, the DST evidence can be useful at a higher level of abstraction. In situations in which there is sufficient information to assign probabilities to individual matches, the Dempster-Shafer theory can be simplified into a more conventional form of probabilistic thinking.

The unpredictability of system responses can be explicitly depicted when an imprecise input is characterized by a collection, or a period and the consequent output is similarly a set or an interval. This is the case when both the input and the resultant output have the same characteristics.

Research Question(s)

To implement the Dempster-Shafer Theory algorithm as well as check the prediction accuracy using this mentioned algorithm in the iris dataset. The actual aim of this paper is to understand and implement the Dempster-Shafer Theory using the python programming language. To reflect doubt and inaccuracies in the evidence, one might use a belief structure that is defined as a set function m that satisfies. The conjecture of this report is to get the closable prediction accuracy that can help to understand future improvement as well as the gap of the applied implementation techniques (Nachappa et al, 2020).

Materials and Methods

D-S theory has been proven to be an effective combination tool in the past, although the majority of research so far has focused on bringing together the outcomes of multiple independent classification strategies. For instance, in a damage detection situation with induction motors, the D-S theory is employed in conjunction with such a neural network methodology. The DRC is a data fusion technique that first uses the neural network to categorize eight different fault states before converting that data into a bulk function assignment. By combining them, DRC helps bring diagnostic certainty closer. They claim that now the D-S methodology's strength comes from its capacity to effectively mix information measurements from different classifiers. In other words, the shortcomings of separate classifiers become far less noticeable when their findings are pooled. As I said before, this study is distinct from the aforementioned methods in that it is solely concerned with classification utilizing the D-S theory. This represents a fresh take on the D-S concept as a unifying classification framework, as previous research has mostly concentrated on tweaking the theory's approach. Use a prediction average accuracy for the mass equations to demonstrate a novel approach to applying D-S for monitoring systems as well as problem diagnostics. By omitting the need to deal with the assignment problem of opposing beliefs, the authors believe their architecture outperforms more conventional mass assignment methods (Fei et al, 2019). In addition, this paper also worked with the iris dataset to get a better understanding as well as get a proper view of the DempsterShafer Theory with a dataset. The common UCI datasets benchmark is the Iris Flower Dataset [IPD]. Many researchers use it as a classification testing option. The collection consists of 150 unique records, each of which is accompanied by four numerical attributes: sepal length, sepal width, petal length, and petal width (all in cm) (Papasotiriou, 2021). Each of the three groups (Iris Setosa, Iris Versicolor, or Iris Virginica) contains 50 specimens in total.

For the IPD, "Setosa, Versicolor, Virginia" serves as the discerning framework, as well as the initial categorization step provides for seven hypotheses:

- {Setosa} = class 1
- {Versicolour} = class 2
- {Virginica} = class 3
- {Setosa, Versicolour}
- {Setosa, Virginica}
- {Versicolour, Virginica}
- {Setosa, Versicolour, Virginica}

Results

This section helps to understand the Dempster-Shafer implementation with the iris dataset data in step by step. Pick 10 records at random from the IPD DataFrame in order to get feel for the information; the Class column will show you how the data is commonly categorized for Iris (Huang et al, 2021).

```
data_file = r'file_location/iris.csv'
df = pd.read_csv(data_file).dropna()
print('data shape:', df.shape)
df.sample(10)
```

```
data shape: (150, 6)
```

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
108	109	6.7	2.5	5.8	1.8	Iris-virginica
72	73	6.3	2.5	4.9	1.5	Iris-versicolor
1	2	4.9	3.0	1.4	0.2	Iris-setosa
97	98	6.2	2.9	4.3	1.3	Iris-versicolor
126	127	6.2	2.8	4.8	1.8	Iris-virginica
82	83	5.8	2.7	3.9	1.2	Iris-versicolor
85	86	6.0	3.4	4.5	1.6	Iris-versicolor
127	128	6.1	3.0	4.9	1.8	Iris-virginica

Figure 1: Load the iris dataset

The distinct Class elements in the DataFrame df are the source of the IPD's analytical framework.

```
pset = dict.fromkeys(powerset(df.Class.unique()), 0)
print(pset)
```

Figure 2: Specify the parameters of judgment.

The empty set (frozenset()) is represented here as the power setting of the period of discernment.

```
{frozenset(): 0,
frozenset({'Virginica'}): 0,
frozenset({'Setosa', 'Virginica'}): 0,
frozenset({'Versicolor'}): 0,
frozenset({'Versicolor', 'Virginica'}): 0,
frozenset({'Setosa'}): 0,
frozenset({'Setosa', 'Versicolor'}): 0,
frozenset({'Setosa', 'Versicolor', 'Virginica'}): 0}
```

Figure 3: Specify the parameters of judgment outcomes.

Since all IPD attributes are numbers, class members may be quickly and easily determined by looking at the range of values for that attribute. The class members will be determined in the future based on the values of these attributes.

```
classRange = {}
for c in df.Class.unique():
    fieldRange = {}
    for f in fields:
        fieldRange[f] =(df[df.Class == c][f].min(), df[df.Class == c][f].max())
    classRange[c] = fieldRange
print(classRange)
```

Figure 4: Identify a group's membership

We determined the minimum and maximum values (min, max) for each property in the class's information,

```
{
'Setosa': {
'PetalLength': (1.0, 1.9),
'PetalWidth': (0.1, 0.6),
'SepalLength': (4.3, 5.8),
'SepalWidth': (2.3, 4.4)},
'Versicolor': {
'PetalLength': (3.0, 5.1),
'PetalWidth': (1.0, 1.8),
'SepalLength': (4.9, 7.0),
'SepalWidth': (2.0, 3.4)},
'Virginica': {
'PetalLength': (4.5, 6.9),
'PetalWidth': (1.4, 2.5),
'SepalLength': (4.9, 7.9),
'SepalWidth': (2.2, 3.8)}
)}
```

Figure 5: Identify a group's membership outcomes

The estimated accuracy of the predictions is 96%! According to the findings, this D-S classifier performs comparably to other top methods when it comes to sorting items into one of several categories using a limited number of features.

```
total rows 150
total predicted 150
total fail predicted 6
accuracy rate 0.96
```

Figure 6: Accuracy prediction outcomes

Conclusions

This paper concludes the details of the Dempster-Shafer Theory as well as the implementation techniques using the python language. In addition, the methodology section and the literature review section is aims to share brief details of this mentioned topic. On the other hand, the implementation techniques are also included in this paper to get a better understanding as well as get an optimistic view of this mentioned topic (Rousaki, 2020). The future work for this algorithm can be improved using some more advance technical concepts or technology. In addition, the algorithm can be developed with any other and more efficient logic that also helps to get more accurate data as well as other outcomes. Moreover, based on the dataset this mentioned algorithm works and returned almost 96% accuracy so if the dataset also improved or adds more data about the iris flower then the result also returned more accurate.

Acknowledgments

- Dan Lo

Contact Information

1.snannap2@students.kennesaw.edu
2.dbandi@students.kennesaw.edu
3.tyarlaga@students.kennesaw.edu
4.jrayavar@students.kennesaw.edu
5.seadara@students.kennesaw.edu

References

1.Du, Y.W. and Zhong, J.J., 2021. Generalized combination rule for evidential reasoning approach and Dempster–Shafer theory of evidence. Information Sciences, 547, pp.1201-1232.

2.Fei, L., Xia, J., Feng, Y. and Liu, L., 2019. An ELECTRE-based multiple criteria decision making method for supplier selection using Dempster-Shafer theory. IEEE Access, 7, pp.84701-84716.

3.Huang, L., Denœux, T., Tonnelet, D., Decazes, P. and Ruan, S., 2021, September. Deep PET/CT fusion with Dempster-Shafer theory for lymphoma segmentation. In International Workshop on Machine Learning in Medical Imaging (pp. 30-39). Springer, Cham.

4.Kang, B., Zhang, P., Gao, Z., Chhipi-Shrestha, G., Hewage, K. and Sadiq, R., 2020. Environmental assessment under uncertainty using Dempster–Shafer theory and Z-numbers. Journal of Ambient Intelligence and Humanized Computing, 11(5), pp.2041- 2060.

5. Nachappa, T.G., Piralilou, S.T., Gholamnia, K., Ghorbanzadeh, O., Rahmati, O. and Blaschke, T., 2020. Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. Journal of hydrology, 590, p.125275.

6.Papasotiriou, I.B., 2021. Movie Recommendation System based on Dempster-Shafer theory.

7.Rousaki, A., 2020. Music Recommendation Systems using Dempster Shafer Theory (Doctoral dissertation, BSc Thesis, Departmnet of Informatics and Telecommunications, NKUA).

8.Ullah, I., Youn, J. and Han, Y.H., 2021. Multisensor data fusion based on modified belief entropy in Dempster–Shafer theory for smart environment. IEEE Access,

9, pp.37813-37822. 9.Xiao, F., 2020. Generalization of Dempster–Shafer theory: A complex mass function. Applied Intelligence, 50(10), pp.3266- 3275.

KENNESAW STATE UNIVERSITY
COLLEGE OF COMPUTING AND SOFTWARE ENGINEERING

Author(s): Sri Manju Nannapaneni , Dhanunjai Bandi, Taraka Vishnu Yarlagadda, Jayanth Kumar Rayavarapu, Sandeep Eadara

Advisors(s): Dan Lo