Chap4#5#2: Machine Learning for Anomaly-based Spam Detection

April 12, 2023

भारतीय प्रौद्योगिकी
संस्थान जम्मू
INDIAN INSTITUTE OF
TECHNOLOGY JAMMU

Devesh C Jinwala,
Professor, SVNIT and Adjunct Prof., CSE, IIT Jammu
Department of Computer Science and Engineering,
Sardar Vallabhhai National Institute of Technology, SURAT

Topics to study in Chapter 4

- Machine learning for Anomaly Detection: Definition of an anomaly. Types of Anomalies or outliers in machine learning. Motivation for machine learning for anomaly detection.
  Data Visualization. Supervised, Unsupervised and Semi-supervised Learning methods for Anomaly Detection.
  Applications of Anomaly Detection: Intrusion detection, Fraud detection, Health monitoring, Defect detection, and lastly Spam detection. Intrusion Detection with Heuristics.Goodness-of-fit. Host Intrusion Detection. Network Intrusion Detection. Web Application Intrusion Detection.
  Overview of Machine learning Approaches for Anomaly Detection: Distance-based, Clustering-based and Model-based Approaches. Algorithms for Distance and Density-based approaches, Rank-based approaches, Ensemble Methods Algorithms for Time Series Data. Deep Learning for Anomaly Detection. Behavioural-based Anomaly Detection [8 hours]

Topics in Handouts#1, #2, #3

1. Email Basics, Spam Detection Basics, and ML-based Spam Detection Basics
2. Classical ML-based Spam Filtering
3. Deep Learning-based Spam Filtering

ML-based Methods used for Email Spam filtering

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- is a type of approach used in dividing objects or case examinations into comparatively similar collections known as clusters.

[1] J.S. Whissell, C.L.A. Clarke, Clustering for semi-supervised spam filtering, In: Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS '11), 2011, pp. 125–134.

[2] Spam e-mails filtering techniques. In: Int. J. Tech. Res. Appl., 4 (6), pp. 7 - 11, 2016

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- is a type of approach used in dividing objects or case examinations into comparatively similar collections known as clusters.
- algorithms are unsupervised learning tools are used on e-mail Spam datasets which usually have true labels.

[1] J.S. Whissell, C.L.A. Clarke, Clustering for semi-supervised spam filtering, In: Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS '11), 2011, pp. 125–134.

[2] Spam e-mails filtering techniques. In: Int. J. Tech. Res. Appl., 4 (6), pp. 7 - 11, 2016

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- is a type of approach used in dividing objects or case examinations into comparatively similar collections known as clusters.
- algorithms are unsupervised learning tools are used on e-mail Spam datasets which usually have true labels.
- in research a good number of clustering algorithms have been shown to be effective i.e. to classify e-mail Spam datasets into either ham or Spam clusters (e.g. in Whissell and Clarke in[1]),

[1] J.S. Whissell, C.L.A. Clarke, Clustering for semi-supervised spam filtering, In: Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS '11), 2011, pp. 125–134.

[2] Spam e-mails filtering techniques. In: Int. J. Tech. Res. Appl., 4 (6), pp. 7 - 11, 2016

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- is a type of approach used in dividing objects or case examinations into comparatively similar collections known as clusters.
- algorithms are unsupervised learning tools are used on e-mail Spam datasets which usually have true labels.
- in research a good number of clustering algorithms have been shown to be effective i.e. to classify e-mail Spam datasets into either ham or Spam clusters (e.g. in Whissell and Clarke in[1]),
- Two types of clustering methods that have been used for Spam classification as such.

[1] J.S. Whissell, C.L.A. Clarke, Clustering for semi-supervised spam filtering, In: Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS '11), 2011, pp. 125–134.

[2] Spam e-mails filtering techniques. In: Int. J. Tech. Res. Appl., 4 (6), pp. 7 - 11, 2016

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- is a type of approach used in dividing objects or case examinations into comparatively similar collections known as clusters.
- algorithms are unsupervised learning tools are used on e-mail Spam datasets which usually have true labels.
- in research a good number of clustering algorithms have been shown to be effective i.e. to classify e-mail Spam datasets into either ham or Spam clusters (e.g. in Whissell and Clarke in[1]),
- Two types of clustering methods that have been used for Spam classification as such.
- Density-based clustering and K-nearest neighbours (kNN).

[1] J.S. Whissell, C.L.A. Clarke, Clustering for semi-supervised spam filtering, In: Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS '11), 2011, pp. 125–134.

[2] Spam e-mails filtering techniques. In: Int. J. Tech. Res. Appl., 4 (6), pp. 7 - 11, 2016

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- is a type of approach used in dividing objects or case examinations into comparatively similar collections known as clusters.
- algorithms are unsupervised learning tools are used on e-mail Spam datasets which usually have true labels.
- in research a good number of clustering algorithms have been shown to be effective i.e. to classify e-mail Spam datasets into either ham or Spam clusters (e.g. in Whissell and Clarke in[1]),
- Two types of clustering methods that have been used for Spam classification as such.
- Density-based clustering and K-nearest neighbours (kNN).
  - density based clustering implemented in[2] showed the capacity to process encrypted messages too, thereby upholding privacy confidentiality.

[1] J.S. Whissell, C.L.A. Clarke, Clustering for semi-supervised spam filtering, In: Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS '11), 2011, pp. 125–134.

[2] Spam e-mails filtering techniques. In: Int. J. Tech. Res. Appl., 4 (6), pp. 7 - 11, 2016

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- kNN proposed in [1] is a distribution free method, - i.e. the data is not required to be drawn from a given probability distribution.

[1] https://saravananthirumuruganatha n.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighborknn-algorithm/.

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- kNN proposed in [1] is a distribution free method, - i.e. the data is not required to be drawn from a given probability distribution.
    - this property is very vital as in the actual scenario, nearly all of the applied data disobey the standard hypothetical postulations made (such as Gaussian mixture, linearly separable, and others).

[1] https://saravananthirumuruganatha n.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighborknn-algorithm/.

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- kNN proposed in [1] is a distribution free method, - i.e. the data is not required to be drawn from a given probability distribution.
  - this property is very vital as in the actual scenario, nearly all of the applied data disobey the standard hypothetical postulations made (such as Gaussian mixture, linearly separable, and others).
- here, the classification model is not built from data, rather classification is carried out by matching the test instance with K training examples and decision is made as to which group it belong to depending on the resemblance to K closest neighbors.

[1] https://saravananthirumuruganatha n.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighborknn-algorithm/.

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- kNN proposed in [1] is a distribution free method, - i.e. the data is not required to be drawn from a given probability distribution.
  - this property is very vital as in the actual scenario, nearly all of the applied data disobey the standard hypothetical postulations made (such as Gaussian mixture, linearly separable, and others).
- here, the classification model is not built from data, rather classification is carried out by matching the test instance with K training examples and decision is made as to which group it belong to depending on the resemblance to K closest neighbors.
- is termed as a lazy learner since the training data points is not used by it to perform generalization. That is, there is no obvious training stage and if it exists it is extremely small.

[1] https://saravananthirumuruganatha n.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighborknn-algorithm/.

ML-based Spam filtering: Clustering

Clustering-based ML Spam filtering. Clustering....

- kNN proposed in [1] is a distribution free method, - i.e. the data is not required to be drawn from a given probability distribution.
    - this property is very vital as in the actual scenario, nearly all of the applied data disobey the standard hypothetical postulations made (such as Gaussian mixture, linearly separable, and others).
- here, the classification model is not built from data, rather classification is carried out by matching the test instance with K training examples and decision is made as to which group it belong to depending on the resemblance to K closest neighbors.
- is termed as a lazy learner since the training data points is not used by it to perform generalization. That is, there is no obvious training stage and if it exists it is extremely small.
- the implication is that the algorithm has a moderately speedy training phase.

[1] https://saravananthirumuruganatha n.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighborknn-algorithm/.

ML-based Spam filtering: Clustering

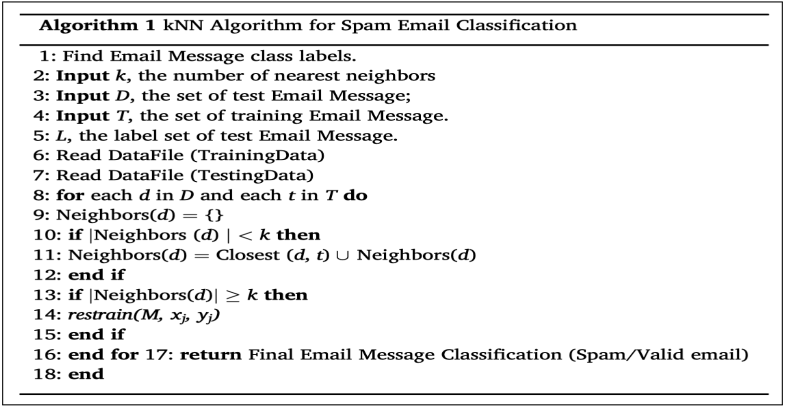
Clustering-based ML Spam filtering. Clustering....

- kNN proposed in [1] is a distribution free method, - i.e. the data is not required to be drawn from a given probability distribution.
  - this property is very vital as in the actual scenario, nearly all of the applied data disobey the standard hypothetical postulations made (such as Gaussian mixture, linearly separable, and others).
- here, the classification model is not built from data, rather classification is carried out by matching the test instance with K training examples and decision is made as to which group it belong to depending on the resemblance to K closest neighbors.
- is termed as a lazy learner since the training data points is not used by it to perform generalization. That is, there is no obvious training stage and if it exists it is extremely small.
- the implication is that the algorithm has a moderately speedy training phase.
- however, the entire training data is required throughout the testing phase as decisions are made based on the complete training data set.

[1] https://saravananthirumuruganatha n.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighborknn-algorithm/.

ML-based Spam filtering: Clustering: kNN algorithm pseudocode

Here Neighbours(d) return the k nearest neighbours of d, Closest (d, t) return the closest elements of t in d, and testClass(S) return the class label of S.



Algorithm 1 kNN Algorithm for Spam Email Classification

```
Find Email Message class labels.
Input $k$, the number of nearest neighbors
Input $D$, the set of test Email Message;
Input $T$, the set of training Email Message.
$L$, the label set of test Email Message.
Read DataFile (TrainingData)
Read DataFile (TestingData)
for each $d$ in $D$ and each $t$ in $T$ do
    Neighbors($d$) = {}
    if |Neighbors ($d$) | $<$ $k$ then
        Neighbors($d$) = Closest ($d$, $t$) $\cup$ Neighbors($d$)
    end if
    if |Neighbors($d$)| $\geq$ $k$ then
        restrain($M$, $x_j$, $y_j$)
    end if
end for return Final Email Message Classification (Spam/Valid email)
end
```

Figure: KNN based Clustering for Spam detection[1]

[1] S. Zhu, W. Dong, W. Liu, Hierarchical reinforcement learning based on KNN classification algorithms, Int. J. Hosp. Inf. Technol. 8 (8) (2015) 175–184.