FAKE NEWS DETECTION USING NLP

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE AWARD OF

BACHELOR OF ENGINEERING DEGREE IN COMPUTER SCIENCE AND ENGINEERING

BY

EZHILKAVIYA.K

DEPARTMENT OF COMPUTER SCEINCE AND ENGINEERING

STARLION COLLEGE OF ENGINEERING AND TECHNOLOGY

MANANKORAI,THANJAVUR-614206

OCTOBER-2023

OBJECT DETECTION WITH YOLO

Abstract

The field of artificial intelligence is built on object detection techniques. YOU ONLY LOOK ONCE (YOLO) algorithm and it's more evolved versions are briefly described in this research survey. This survey is all about YOLO and convolution neural networks (CNN) in the direction of real time object detection. YOLO does generalized object representation more effectively without precision losses than other object detection models. CNN architecture models have the ability to eliminate highlights and identify objects in any given image. When implemented appropriately, CNN models can address issues like deformity diagnosis, creating educational or instructive application, etc. This article reached at number of observations and perspective findings through the analysis. Also it provides support for the focused visual information and feature extraction in the financial and other industries, highlights the method of target detection and feature selection, and briefly describe the development process of yolo algorithm.

Introduction

Object detection strategies are the establishment for the engineered insights subject. This paper offers a brief evaluation of the you only look once (YOLO) algorithm of rules and another prevalent variation. Through the assessment, it is understood that numerous comments and proper results show the variations and likenesses between Yolo version and between Yolo and convolutional neural networks (CNNS). The relevant perception is the Yolo algorithm development continues to be Ongoing. This article in brief depicts the improvement procedure of the Yolo set of rules, summarizes the methods of goal Recognition and characteristic choice, and provides literature assistance for the centered image news and characteristic extraction YOLO is a new approach of object detection. The classifiers are used to detection in earlier works in an object detection. It is consider here that frame object detection as a regression issue to spatially separated bounding package container and associated class probabilities as an alternative object detection Real Time Object Detection System with YOLO and CNN Models.

A Review In a single evaluation, a single neural network predicts a boundary bins and class possibilities directly from complete images. The entire direction pipeline may be optimized start to finish because it is a single network. On the effectiveness of detection without delay a unified architecture moves incredibly quickly and as well 45 frames per 2D frame yolo model process image in real time. Yolo, a scaled down version of the network, achieves double the coverage of the other real time detectors while operating at an incredible 125 frames per second. Yolo makes more localization errors than modern detection algorithm but is less likely to anticipate false positives based on past data. Yolo eventually picks up and highly preferred representation of items while generalizing from natural photographs to other domain names like painting, its outperform competing detection techniques like DPM and R-CNN .

Approaches of implementation of YOLO and CNN

Object detection is a technology that detects the semantic Objects of a category in virtual snap shots and films. Certainly, one of its Actual-time packages is self-riding automobiles. In this, our challenge is to locate a couple of items from a photo. The maximum common Item to come across on this utility is the car, motorcycle, and Pedestrian. For locating the gadgets within the photograph, Object localization and should find multiple item in real-time structures. There are various techniques for item Detection, they can be break up into classes, first is the Algorithms primarily based on classifications. CNN and RNN come below this category. On this, pick out the involved Regions from the picture and ought to classify them the use of Convolutional neural community. This technique may be very sluggish Due to the fact it should run a prediction for every decided on Vicinity. The second one class is the algorithms primarily based on Regressions. Yolo approach comes below this category. In This, it might not choose the fascinated regions from the photograph. Instead, it expect the training and bounding containers of the complete picture at a single run of the algorithm and detect a couple of gadgets using an unmarried neural community. Yolo Set of rules is rapid as compared to other classification Algorithms. In actual time our algorithm technique 45 frames consistent with 2d. Yolo algorithm makes localization mistakes however Predicts less fake positives in the background. Humans can without problems come across and pick out gadgets of their environment without attention in their instances, irrespective of what position they're in and whether or not they are the wrong way up, one-of-a-kind in shade or texture, in part occluded, and so forth. To extract information about the objects and shapes in a picture, the same item detection and recognition on a computer requires a lot of processing. Identifying an object in an image or video is referred to end as an object detection in computer vision. The main steps in object detection or future extraction, which is imported for surveillance, cancer reduction, car identification, and underwater object detection, among other applications. Different approaches had been used to accurately and correctly identify the object for specialized packages. These suggested solutions still have an issue with inaccuracy and inefficiency, though. Device learning and deep neural network approaches are more successful in addressing these issues of object detection. Convolution neural networks CNN for widely used in image processing and other areas with the development of artificial intelligence because of their fantastic performance. However, because it's a set of rules is so complex computationally, CNN faces enormous difficulties capturing the interest of mobile devices. FPGA becoming ideal choices for CNN deployment due to their advantages of high performance, reprogramming, and seldom energy usage. The yolo algorithm views target detection as a regression problem in comparison to other CNN algorithms.

It is a one-step algorithm that runs quickly and requires little calculation. It is much more appropriate for FPGA hardware structure. The issue of massive computational of CNN and constrained assets on FPGA chip is resolved, and the parallelism of FPGA is used to boost the CNN. This is accomplished by improving the Yolo network version and fixed point, among other things. Experimental finding demonstrate that the strategy suggested in this paper as significantly reduced the operating cost while maintaining accuracy and as an essential reasonable cost within the creation of mobile terminals and real time computing. In order to speed up the deep learning yolo community, this paper offers an advanced set of guidelines for deep learning Yolo network that are entirely based on Xilinx FPGA.

By using the parallelism function of the FPGA to speed up the CNN, the acceleration set of rules specifically addresses the issue that the CNN requires a significant amount of calculation while the FPGA has limited resources on the chip. The implementation of this set of rules specifically includes all three parts: the first one is the FPGA Yolo community design the second one is the selection of the activation characteristics and the last one is 16 bit constant point optimization of the weight parameter. Development plan is presented as a result of the lack of common accuracy and miss detection in the way of actual roots in gold detection through the Yolo V3 community in order to investigate the anchor range and component ratio, the original network clustering set of rules is updated using the ok-means clustering algorithm similarly, so that you can enhance the performance of the road target detection set of rules, the present network output is upgraded, and a 104*104 feature detection layer is brought, and the feature map output through 8 instances sampling can be output through 2 instances up sampling, and 4 the characteristic maps of down-sampling are stitched together, and the 104*104 feature maps received can efficiently lessen the disappearance of features. Via the experimental effects, it is able to see that compared with the stepped forward yolov3 set of rules, the common detection accuracy of the stepped forward set of rules for road goal detection is quite excessive to a few.17%, and the missing detection rate is reduced by 5.62%. In this paper author proposed laptop vision functions of onboard cameras and embedded devices are widely employed with general unmanned aerial vehicles (UAV). However ,it is very difficult program to study the real time seen through the object identification approach go to the limited memory and computing capability of embedded devices at the UAV platform. This study evaluates the performance of different you yolo collection methods on the Pascal voc dataset, employees map and frame per second at assessment measures, and applies the learning outcomes to the xtdrone UAV simulation platform for testing in order to overcome those difficulties. The map of Yolo V4 is 87.48%, which is 14.2% better than that of yolo V3 according to our evaluation of Yolo V3, Yolo V3 tiny, and yoov3-spp3 and yolov4-tiny on the pascal voc benchmark dataset. Frame per second reaches 72, and the test sets check time is one or 103.8 seconds, the validation sets shortest test time is Yolo V4, but the map only achieves 50.06% accuracy which is less precise than Yolo V3 tiny. Using simulation on the xtdrone platform to analyze the performance of five modules on the Pascal voc data set come on this paper ultimately comes to the conclusion that Yolo V3 can match the demands of real time, lightweight, and excessive precision. The main objective of real time object detection is to locate the location of an object in a supply picture accurately and mark the item with the appropriate category. In this paper it used actual time object detection you appearance best as soon as (yolo) set of rules to train our device studying version. Yolo is a clever neural network for doing object detection in actual time and with the assist of coco dataset the set of rules is trained to perceive one of a kind items in a specific picture. After training this method locate the object in actual time with ninety% accuracy. The aim of item detection is to come across all times of gadgets from an acknowledged magnificence, consisting of humans, motors or faces in an image. Commonly, most effective a small quantity of instances of the object are gift inside the picture, however there are a very huge range of possible places and scales at which they could occur and that want to by some means be explored. Each of the detection is mentioned with some form of pose data. This could be as simple because the region of the object, an area and scale, or the volume of the object described in terms of a bounding field. In different conditions, the pose facts is extra specific and includes the parameters of a linear or nonlinear transformation. As an example, a face detector may additionally compute the places of the eyes, nose and mouth, further to the bounding box of the face. Synthetic intelligence is being tailored by the sector on the grounds that past few years and deep gaining knowledge of performed a crucial position in it. This paper focuses on in depth learning and how it is used to find and tune the devices. Deep learning uses algorithms that are impacted by the structure and functions of the brain. Working with such algorithms as the advantage that performance generally improves with growth in data, as opposed to conventional learning algorithms whose performance stabilizes regardless of boom in the amount of data a number of most popular research topics in computer vision application have put real time item monitoring at the forefront. Despite the impressive progress made at this paper, it is still difficult to effectively and accurately track the objects in real time at a significant level. The capabilities of images and movies for security and oversight packages are extracted while defining detection and monitoring algorithm. You only look once (Yolo), location based fully convolutional neural network (RCNN), and quicker samples of popular object detection algorithm (f- RCNN). RCNN is more accurate than other algorithms, but yolo outperforms them when speed is prioritized over accuracy. Yolo treats object detection as a regression issue and provides elegance chances for found images.

Real world circumstances drive the need to use the Yolo algorithm to find object entities base. When pace is taken into account as if performance indicator, Yolo outperforms other object detection algorithms. Regression is used by the community to predict the bounding containers and classes of entities in a single pass of algorithm, as opposed traversing the network more than once to identify and output the opportunity of the elegance. This gives it a real and noticeable advantages over earlier technology in terms of increased frame per second. When compared to other performances like CNN and its various version this result in an object detector using Yolo V3 with impressive overall performance metrics, including accuracy, precision, and everything it is different from other. This enables the builders to use the Yolo V3 for real time monitoring packages in order to benefit from its performance improvements over its antiquated counterparts. By using photographic records enhancement, it is possible to expand the pool of records without actually adding any new information. In this study, researchers assessed the effectiveness of image manipulation techniques like the double phantasm and inversion on the overall performance of the face detection for information enhancement needs. The study found that the limited amount of data available to create a teaching version of the facts was a weakness in the facts that were acquired. The goal of this people like research paper is to increase the variety of the facts such that come up and gives various comparable datasets, it can make prediction with accuracy. Yolo V3 is used to perform face detection on photos, and accuracy results from the dataset and before including records are compared. Object detection as received a lot of attention in the field of computer vision carried out in various areas, including the clinical field, pastime reputation, security tracking in agencies, and automated manage robots. Famous item detection techniques, before everything, had been recognized by way of the use of feature extraction techniques together with histogram of orientated gradients (hog), speeded-up strong features (surf), nearby binary styles (lbp), and additionally coloration histograms. The system of taking pictures the characteristics of gadgets that could describe the traits of items is the number one approach in the function extraction technique. A straightforward, ready-to-use, unified object detection model that works with complete images. YOLO also generalized well to unused spaces utilized in applications that depend on quick, strong object detection. A degenerative show built for recognizing degraded images like obscured and loud pictures has the model being prepared with these debased pictures. This model performed superior in terms of detecting degraded pictures and adapted way better with complex scenes. For discovery shallow person on foot highlights, a YOLO v2 organize was adjusted by including three Pass through layer to them. The number of detection frames can reach 25 frames per second, which meets the demands of real-time performance. To recognize indoor impediments an unused strategy of using profound learning together with a light field camera was utilized. The strategy recognizes the deterrents and perceives its information. YOLO connected to vehicle entryway board welding panel lines can distinguish and distinguish patch joint accurately the calculation can too identify the position of the solder joints and more. Object Detection has received a lot of research attention in recent years because of its tight association with video analysis and picture interpretation. The foundation of conventional object detection technique is shallow trainable structures and handmade features. Building intricate ensembles that incorporate several low-level picture features with high-level context from object detectors and scene classifier can readily stabilize their performance. In order to solve the issues with traditional architectures, more potent tools that can learn semantic, high-level and deeper features are being offered as a result of deep learning's quick development. In terms of network architecture, training methodology, optimization function, etc., these models behave differently. In this paper, it explore object detection frameworks based on deep learning. A brief background of deep learning is provided before our review. It brief history of deep learning and its illustrative tool, the convolutional neural network (CNN), is given before our review. Then it concentrate on common generic object detection architectures with a few changes and helpful tips to further enhance detection performance. This paper also provides overview of a number of specific task, such as salient item detection, face detection and pedestal detection, as different specific detection task exhibit different characteristics. To evaluate different approaches and get some insightful results experimental analysis are also offered. In order to provide direction for future work object identification and pertinent neural network based learning systems, a number of promising directions and tasks are provided. An aspect of your computer vision that is constantly in use is object detection. Object identification has made remarkable progress, from the first manual future extractor to the future extractor dominated by deep convolution network. Researchers started paying more attention to CNN in 2012 when Alexnet won the title image net with significant benefits (convolutional neural report). CNN has made remarkable success since 2012 in many facets of computer science, and object vision detection is no exception. As of now, two step end single step CNN-based object detection network can be loosely split into two categories. The two stage object detection network typically consists of three stages: The first stage involves creating multiple delimitations for the proposal cans; the second stage involves performing the object classification. R CNN, Fast R-CNN, SPPNet,R-CNN faster, FPN, and Mask- RCNN are some of the representative network for the prediction of position information on the bounding boxes of the proposition. The single phase object detection network is an end to end model, meaning that only one network inference is required from the input image to the network output. Its representative objectives include the YOLO series, SSD and RetinaNet. The benefits and drawbacks of the two strategies mentioned above very. Although the single stage detector is faster than the two stage one, the location accuracy of the two stage detector is marginally higher. However, there is a significant issue for practical application computationally expensive and memory consuming. This problem exists independent of the type of detecting network. In the field of goal recognition, YOLO algorithm has carried out well. In this paper, it enhance the brand new YOLO community version via way of means of converting the residual devices to dense connection in the CSP module and including channel interest mechanism. The advanced community version alleviates the vanishing- gradient hassle, complements function propagation, encourages function reuse, and decreases the quantity of parameters. What's more, it can adaptively recalibrate the channel records of the function maps and enhance the overall performance of goal detection. Experimental consequences display that the advanced YOLO community version greatly improves the detection accuracy. In addition, it optimizes the hassle of lacking and miss-detecting targets convolutional neural network (CNN) has end up the famous deep gaining knowledge of set of rules for goal recognition. However, while the function maps attain the stop of the community structure after a couple of convolutions, the function facts in it may disappear. To clear up this type of problem, pupils placed forward a few answers and fashions, consisting of ResNet and DenseNet . Each of those models has a key function that they devise a short route among the front community layer and the again community layer. In the ResNet community model, channels are brought up, but channels are linked within the DenseNet community model. It guarantees that the facts go with the drift among layers within the community is maximized. And as it does now no longer want to analyze redundant function maps, every layer community can attain extra function facts with fewer parameters. In addition, DenseNet's facts go with the drift and gradients throughout the community are improved, in order that it turns into less difficult to be trained. Convolution neural networks [CNN] were widely used in photographs and other parts as artificial intelligence (AI) advanced due to their superior performance. Although CNN faces significant challenges in the development of mobile devices due to its computationally complex set of algorithms. FPGA are good candidates for CNN deployment because of their high performance, re programmability, and sporadic energy usage. The YOLO set of rules views goal detection as a regression issue in comparison to various CNN algorithms. It is a one-step set of rules that execute quickly and require little calculation it is suitable for hardware platform to implement. This paper suggests a sophisticated set of guidelines for a deep learning YOLO community that are entirely based on Xilinx ZYNQ FPGA. The issue of large computational of CNN and confined assets on FPGA chips as resolved by improving the YOLO community version and fixed point, etc. The parallelism FPGA is then exploited to increase the seat and length. Results from experiments show that the method suggested in this study as greatly improved operational cost while maintaining accuracy and as a very reasonable cost inside the consideration of CNN real time computing. In this paper, the author proposed a method to classify human and automotive objects by integrating a support vector machine (SVM) with the deep learning version, you only look once (Yolo), in a high resolution automotive radar device. The constraints of gold envision from the yellow version or incorporated to the SVM in order to improve the overall performance of the class. The result from the yolo and SVM may then be combined with the predetermined target bonds to advance the overall class accuracy. The results showed that they suggested method outperforms those obtained using yolo or in terms of overall performance. A new device version that combines the conventional method, an SVM, and the deep learning version, you only look once (yolo), to categorize people and motors. Based on real measurements, the proposed technique can classify people and motors with over 90% accuracy in high decision automotive radar will stop using the suggested version, the SVM and yolo version are training using range angle (RA) domain information from preprocessed radar signal to classify the same scene. In particular, the enhanced class performance by utilizing boundary containers from the yolo version. Finally, by integrating every result, the overall class accuracy is increased. In this paper the author describes the CPU based yolo, a real time object detection version that may be used with non GPU computers to benefit users of low configuration computers. Many well developed algorithms are available for object detection including yolo, Faster R-CNN, Fast R-CNN R, R-CNN, mask R-CNN, R-FCN, SSD, Retinanet, and others. YOLO is a set of deep neural network rules for object detection that is more accurate and quick than most other algorithms. YOLO is made of four computers that are entirely GPU based and media graphics card with at least 12 GB of memory. In this paper they are implemented yolo with open CV in order to make real time object detection possible on CPU based computers. On several non GPU computers, our version accurately and reliably detects an object from video at a frame rate between 10.12 and 6.29 frame per second. The mAP for CPU based Yolo is 31.5% the evaluation of computer vision using deep learning as being built up and completed throughout time mostly using one specific, well known set of rules called convolution neural networks. It has a convolution layer where the output is produced by convolving a clear output with a variety of input components. The environment of a convolution layer allows for the extraction of related styles from companion.

Further more, compared to a tightly linked layer, a convolution layer tends to require a considerably less weight to be determined because filter are not used in convolution layer.

This study builds a smart monitoring device in response to the current status of the monitoring device, which requires human management and isn't intelligent enough. The device completes the evaluation of the pictures content, implement smart monitoring, and uses deep learning to distinguish things. The program of the smart monitoring gadget is realized using the twisting moment as an example. The yolo V2 version curling automated snapping images device detection record set throughout the test had a precision rate of 94.3%, and the detection speed was 13 frame per second. The yolo version can accommodate the curling the photo taking devices real time requirements. That equipment completed its role of monitoring during the actual photographic process. For the first time, it is applied in this paper to curling gold detection underneath the curling smart monitoring equipment the digital camera platform is operated at the same time based on Yolo V2 curling detection result. The smart monitoring of the curling is completed and combined with this straightforward recording of the curling track. Study shows that yolo V2 fully satisfies the real time requirements with detection accuracy for curling goals of 94.3% and detection rate of 30 frames per second. The yolo based fully curling smart monitoring equipment can guarantee accuracy and real time performance at the same time as well as complete the actual photographic task. The content smart assessment model of the digital camera previews is introduced by a device architecture of the curling smart monitoring device. The teaching of Yolo apply to the smart evaluation version is introduced in the third element along with the evaluation of outcomes. Real-time detection of the pavement surroundings is an essential if part of independent riding technology. This paper gives a real time automobile detection device primarily based totally on embedded devices. Based on the prevailing YOLOv3-tiny neural community shape, this paper proposes a brand new neural community shape the YOLOv3- tiny, and quantify the community parameter in the community. Trading complexity of competing in embedded devices, making the proposal neural community shape more appropriate for embedded devices. The new shape is tested, earlier than quantization the YOLO V3 live's detection precision way of means of the convolution operation and dispatched to the Yolo layer for goal detection. At the same time, the 13x13 function map is merged with the 26x26 function map is dispatched second one yolo layer for goal detection. The entire network shape makes use of yolo layers to hit upon exceptional size of function graphics, in order that the community version can hit upon last targets in addition to small targets. The community shape of YOLO V3 tiny can gain 87.79% mAp at the take a look at set, after quantization of community parameters can gain 69.79% mAP and detection pays can gain 28 FPS. The layer shape of Yolo V3 that is specifically composed of convolution layer and pool layer. A small function map of 13x13 is received via way of means of the usage of pooling layer to carry out 5 down sampling operation at the input picture, after which the function is extracted.

Breed type is type of detected animal with the aid of using the usage of the pc imaginative and prescient method. This observe is accomplished to assess the overall performance of livestock detection and breed type consistent with particular areas at the frame of livestock that differ consistent with their breed. You Only Look Once (YOLO) the set of rules is used for livestock detection and breed type on the information set. A unique information set is generated using images gathered from Google images in order to demonstrate the usefulness of the suggested strategy in the detection and type of livestock. Experimental results show that the YOLO set of rules can accurately identify the breed of animals on autograph with 92.85% accuracy. With the use of item detection set of rules YOLOv4, this observation seeks to livestock dictation and type. The item detection set of criteria, however, provides the detection and type of the cattle in the photograph. YOLO are used to identify livestock and categories breed for this purpose. The data set was manually constructed from images collected from various structure because there may not be a formally educated version and information sector version education, the title said was manually generated using photographs collected from various structures. As it expand this data set, it pay particular attention to the gender of animals like calf, cow, and steer. The data set was manually classified after that. The positioning and sophistication of the creatures inside the image are determined as a result of being this tagging. Later, version education is accomplished with the aid of using feeding the version with the pc imaginative and prescient method used in this observe. The experimental outcomes display that the educated version finished 92.85 accuracy at the given take a look at set .

Conclusion

In addition to providing an overview of the real time object detection technique used by YOLO, this paper discusses the core CNN algorithm structure. CNN architecture models have the ability to eliminate highlights and identify objects any given image. When implemented appropriately, CNN models can address issues like deformity diagnosis, creating educational or instructive application, etc. In practice, yolo has a lot of advantages over other CNN algorithm. Yolo can train the complete model in parallel since it is a unified object detection model that is easy to build and train in accordance with its simple loss function. The best speed and accuracy tradeoff for object detection is offered by second major version of YOLO known as YOLOv4. In addition, Yolo is the most advanced object identification technique and is advised for real time object detection since it generalized object representation more effectively than other object detection models. With these accomplishments, it is clear that field of object detection as a bright future.

RECURRENT NEURAL NETWORK

Abstract

After the advent of Web, the number of people who abandoned traditional media channels and started receiving news only through social media has increased. However, this caused an increase of the spread of fake news due to the ease of sharing information. The consequences are various, with one of the main ones being the possible attempts to manipulate public opinion for elections or promotion of movements that can damage rule of law or the institutions that represent it. The objective of this work is to perform fake news detection using Distributed Representations and Recurrent Neural Networks (RNNs). Although fake news detection using RNNs has been already explored in the literature, there is little research on the processing of texts in Portuguese language, which is the focus of this work. For this purpose, distributed representations from texts are generated with three different algorithms (fastText, Glove and word2vec) and used as input features for a Long Short-term Memory Network (LSTM). The approach is evaluated using a publicly available labelled news dataset. The proposed approach shows promising results for all the three distributed representation methods for feature extraction, with the combination word2vec+LSTM providing the best results. The results of the proposed approach shows a better classification performance when compared to simple architectures, while similar results are obtained when the approach is compared to deeper architectures or more complex methods. Index Terms—Fake news detection; Word embedding; Recur- rent neural networks; Long short-term memory

Introduction

Fake news exist even before the emergence of the Internet. Generally, fake news imply that editors use false or misleading information to promote their interests. After the advent of web, more consumers began to abandon traditional media channels, thus producing a growing segment of population receiving news only through social media. The term “fake news” can be applied when information is published with no guarantee of being true, its credibility cannot be proven, or when false information is spread for the purpose of deceiving people. The degree of dissemination of false information that is present in social media has significantly expanded during the recent years, due to the ease of sharing information. Moreover, there is a growing presence of bots1 that manipulate the social media recommendation algorithms by interacting with fake news posts. The main purpose of false news is monetising through advertisements. However, there are claims that even presidential elections in important countries may have been manipulated through the use of false news. Detecting this type of news is not only relevant for the society, but also for the technology and social media companies such as Google, Facebook and Twitter , since it implies the responsibility of the social network (and the company) to avoid mass disinformation to its users. We have seen this emerging concern in combating misinformation during the coronavirus (COVID-19) pandemic. Thus, initiatives have been proposed throughout the years to identify fake news using both automated and manual checking methods. The existence of tools to support automatic detection of fake news in languages other than English, more specifically in Portuguese, is still unsubstantial in both, the industry and academia.

This is due to that the computational methods for this purpose are mostly developed for English language texts. However, some initiatives to detect fake news for the Portuguese language have been proposed in an attempt to reduce their spread. One example is Facebook, which financed the development of chat-bot for the Portuguese language to help users learn how to check the veracity of photos, videos, rumours, news and false statements. Note that the task of identifying fake news among a set of news is a classification problem, which is tackled through algorithms that assign labels to samples, considering a set of characteristics. In the specific context of fake news classification, it is necessary to use text-oriented feature extraction methods and classification algorithms that are able to process the temporal dependencies that appear in sentences and long texts. Regarding text-oriented feature extraction, various Natural Language Processing (NLP) techniques that are able to capture similarity relationships between words have been introduced during the recent years, namely Word2vec, Glove e fast Text. These methods are based on the use of artificial neural networks to obtain dense word representations. On the other hand, considering classification algorithms, Deep Learning (DL) methods have substantially improved the state- of-the-art in several domains, such as speech recognition and object detection. These methods allow computational models composed of several layers of abstraction to learn data representations. A widely used DL method for processing sequential data such as text and speech is the Recurrent Neural Network (RNN), and architectures derived from it such as the Gated Recurrent Unit (GRU) Long Short-Term Memory (LSTM). This work proposes a computational approach for the classification of fake news, written in Portuguese, using a LSTM-type recurrent neural networks and distributed representations. Specifically, this work aims at accomplishing the following objectives: Evaluate the main methods available in the distributed representation literature (word embedding); Compare the classification performance of the chosen word embedding methods; Verify the use of LSTM-type RNNs for text classification, as well as their classification performance for the fake news classification in the Portuguese language; and Compare the proposed approach to other state-of-the-art approaches to text classification. In summary, the main contributions of this work are: experimentation of the main approaches of distributed representations for text classification (word2vec, GloVe and fastText), thus comparing their classification performance; proposal of an approach for the purpose of classification of false news for Portuguese language; Use of LSTM-type

RNNs associated with distributed representation methods for detecting false news, by training our approach using a novel labelled text corpus.

Background

This section presents relevant background aspects for under- standing the method that is proposed in this work. Specifically, this section introduces concepts such as distributed representation and LSTM networks.

A. Word Embedding

Word embedding is a set of techniques that derive from the distributive hypothesis of language, which assumes that the meaning of a word can be inferred from the contexts in which that word is used, and in the context of this work, it is used as distributed representation. Methods based on embeddings represent each word as dense vectors of real values with predefined dimensions. According to , vectors generated from word embedding may be used as input for different tasks of NLP, such as natural language recognition, similarity between words, information retrieval, document classification and sentiment analysis. A word embedding is computed by applying dimensionality reduction to the co occurrence matrix that is produced from a large set of texts, known as corpus. A co-occurrence matrix consists of distributional vectors containing values found in cells of a row.

Therefore, similar terms tend to be positioned in the same neighbourhood region in the vector space, thus making it possible to capture essential language features such as syntax and semantics. In this case, similarity relationships may be inferred from simple measures of distance and similarity between vectors. The most popular word vectorization approaches are Word2vec, Global Vectors (Glove) and fastText .Word2vec: Word2vec is a model proposed by that aims at obtaining ANN-based word vector representations. Word2vec uses an input layer of dimension v, which takes a vector in one-hot representation as input; a hidden layer of dimension n, where n refers to the dimension of the vectors that will represent the words; and an output layer of dimension v, which returns vectors similar to the one-hot representation, but instead of 0's and 1, these vectors express probabilities for each vocabulary word. The one-hot representation is the length of the vocabulary size and it is filled with zeros, except for the index representing the word that is being represented, which is filled with

1)The hidden layer is a fully-connected layer, and its weights are, in fact, the embedding of the word. The output layer generates probabilities for the vocabulary target words. The synapses that connect the input layer to the hidden layer, and the hidden layer to the output layer, can be represented by the matrices $WI_{v\times n}$ and $WO_{n\times v}$, respectively. The operation $r = (x \times WI) \times WO$ represents the computation of the vector x from input layer to hidden layer and then from hidden layer to output layer, the purpose of which is to learn hidden layer weights. Finally, Word2vec is trained using a backpropagation approach and transforms the output layer activation values into probabilities through the softmax function. Word2vec proposes two different training architectures: Continuous Bag-of-Words (CBOW) and Skip gram. In the CBOW architecture, a word w_i is predicted based on its context, which consists of a word window composed of terms before and after w_i . The skip-gram architecture,in turn, seeks to maximise the context prediction based on a word wi given as input. The main difference between the architectures is that in CBOW the input layer is replicated C times, where C is the amount of context words, while in the Skip-gram it is the output layer that goes through this replication.

2) Glove: Glove is a model proposed by, and its difference from the two architectures proposed by Word2vec is that it is a model based on word counts. In Glove, the co- occurrence matrix is used to generate embedding, that is, how often a specific term appears among another in a corpus. The notation used by Glove to calculate the probability of a word j occurring in the context of the word i is defined by the Equation 1.
Pi,j =

xi,j n P

k=0

xi,k

, (1)

where x denotes a vector of co-occurrence, the input xi,j displays the number of times the word j appeared in the context of the word i and n the number of words that co- occur with the word i. In the first instance the embedding of the words are generated randomly, as at the Word2vec. In this sense, proposes a function to perform adjustments in the embedding such that the dot product of their vectors are similar to their values in the co- occurrence table. The function that performs the adjustments is presented in Equation 2.

F(wT i w˜k) = Pik =

Xik Xi

, (2)

where wT i is the transposed vector of wi, w˜k one of the embedding vectors of the word, Pik the ratio of the word i with the context word k and X is the co-occurrence matrix. As the F function is not symmetric, a suitable F0 function is defined in Equation 3, in order to guarantee symmetry.

F0(wi,wk) = wT i w˜j + bi +˜ bi (3) where bi and ˜ bi are the bias for the words wi and wk, respectively. wT i and w˜j are the vectors of embedding. Finally, a summation covering the entire V vocabulary is used to generate the embedding. 3) fastText: fastText is a model developed by a team of researchers from Facebook [6], as an extension of the Word2vec model, where the goal is to learn vectors and perform classification of text. For word representation, fastText is based on the Skip-gram model. The algorithm differs from Word2vec in that it assumes that words are composed of n-grams of their characters. The algorithm learns the vector representations distributed for the n-grams of vocabulary words, and forms the distributed vector representation of a word as the sum of the representations of its n-grams. However, the training process takes longer compared to other models fastText works by sliding a window over the input text and learning all context words from the central word. The difference between fastText and the Skip-gram architecture is in the probability function, where it is changed so that the n-grams are counted.

B. Long Short-Term Memory

LSTM is an RNN architecture designed to model long- term temporal sequences more accurately than conventional RNN. LSTM was introduced by with the motivation of offering better performance by solving the vanishing problem that RNNs naturally suffer when dealing with large data streams. According to, the main difference between LSTM and traditional RNNs is that the LSTM architecture has more parameters and also a strategy of information flow control. This strategy is present in the LSTM hidden layer, and it consists of special units called memory blocks, which contain two basic structures: cells that have connections to other temporal storage cells and weighted units called gates that to control the information flow.

Related work

The work by proposed the classification of fake news extracted from Brazilian websites using text summarization techniques, which consist of reducing large textual parts into smaller ones. In this approach, the word2vec word embedding method with dimensionality of 10 was used for feature extraction. For the classifier, the author proposed the use of two Deep Learning (DL) algorithms and two traditional artificial intelligence algorithms. The DL algorithms used in the work were Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM), while the traditional artificial intelligence algorithms were: Random Forests and Support Vector Machines (SVM). The author reported that the experiments performed with the LSTM neural network achieved a higher accuracy, around 79.3%. Despite getting a result close to 80.00%, it is possible to evaluate that the dimensionality of the embedding vector of the words being 10 can be small, considering that a DL algorithm was used. Another relevant factor for obtaining this result could be the size of the database used in the work, in which 286 news were used for each class. The work proposed by presented predictive models to classify fake news using word embedding techniques combined with Machine Learning (ML) and DL algorithms. The word embedding methods chosen for feature extraction were Glove and word2vec. The version of Glove used in the work was the pre-trained word embedding with 300 dimensional feature vector, trained with corpus. The ML algorithms used in the work were: Logistic Regression3, Random Forests, Naive Bayes, SVM, and xgboost. The DL algorithms were: feed-forward and convolutional neural networks. The best model was the one that combined convolutional neural networks and Glove with an accuracy of a value around 97.50%. The second best accuracy, around 91.37%, was achieved by feed forward neural network combined with word2vec. Although it is not possible to evaluate the Word Embedding techniques because the embeddings of the words were trained using different datasets, it is possible to evaluate that the model with the embeddings trained using another dataset presented the best result. The classification of fake news using the Hierarchical Attention Networks (HAN) architecture and word2vec. HAN allows to visualize the output data through a heat map that gives insight regarding the reason why a certain class is chosen. It highlights the words and sentences that considered most important for the classification by the network. The approach achieved an accuracy value close to 95.35%. On the other hand, proposed a comparative study of word embedding methods for sentiment analysis. The work used a neural network as a classifier and word embeddings for feature extraction. The word embedding methods used in the work were Latent Semantic Anaylsis, word2vec, Glove and Sentiment Specific Word Embedding (SSWE), with feature vectors of dimensions equal to 50, 100 and 300. The objective of a sentiment analysis classification model is to classify the polarity of the review of a movie. The dataset used for training and testing the models is the Internet Movie Database. The results of the experiments showed that the best accuracy was obtained when SSWE was used, achieving an accuracy equal to 82.12%. The work proposed by addresses the classification of fake news using two different classification methods: LSTM and Logistic Regression. The approach using LSTM had as input a word embedding vector obtained through other corpus in Portuguse. On the other hand, the Logistic Regression algorithm used as input a statistical representation. The best result was achieved by the Logistic Regression approach, with an f1-score of 92.8%, whilst the approach based on the use of LSTM and Glove with a dimensionality equal to 1000 obtained a f1-score of 89.6%. The objective of this work is to detect false news using distributed representations and LSTM. For this purpose, the methods fastText, Glove and word2vec are used in order to obtain distributed representations of words. We train and test our approach using a novel labelled text corpus.

Methadology

This section presents the methods proposed to tackle the problem, which is the detection of fake news in texts. We propose methods considering the use of Deep Learning (DL) and text-oriented feature extraction, which have been widely used for text classification. The main hypothesis that we intend to explore in this work addresses the possibility that fake news can be identified through a computational model composed of distributed representations and a kind of Recurrent Neural Networks (RNNs), named Long Short-Term Memory (LSTM) architecture. The LSTM architecture has been selected due to its capability to process long sequences of data and given that it allowed to achieve state-of-art results for various tasks related to text processing. Distributed representations are obtained using techniques known as word embeddings, which allow better generalization for classification models due to similarity relationships between words.

A. Data Preparation

1) Pre-processing: Pre-processing is necessary to standardize the text and select features, removing words that may be considered irrelevant (known as stop-words). First, a word tokenization method is used, which is a method that divides a large sample of text into words. For this, the natural language toolkit NLTK library4 is used. After, as mentioned above, all letters of all words are converted from uppercase to lowercase, and then all graphic accents and diacritics are removed 5, as well as all words considered irrelevant and all non-alphabetic characters. The dataset resulting from the pre-processing stage will be used as input to the word embedding stage (distributed representation) with the three mentioned methods (Word2vec, Glove and fast Text)

2) Distributed representations: After the pre-processing process, it is necessary to generate the distributed representations. As mentioned before, this is achieved using word embedding methods, which allow to generate vectors for each word that will afterwards be used as input to the LSTM model. For this procedure, two different libraries are used: the gensim6 for both Word2vec and fast Text methods, and the GloVe7. The use of gensim (for both Word2vec and fast Text) requires two steps. First, the pre-processed text is loaded and the sentences are converted into a vector, where each element of the vector is a text-word. The second step is the proper training of each model using parameters presented in Table II, with the vector containing the text as input. For generating the word embedding using the Glove library, it is necessary to create a script bash with the parameters shown in Table II, including the directory where the text files are. Finally, the word and its embedding will compose both training and test datasets. Table II shows each parameter necessary to generate the word embedding. Values shown in the Size parameter were chosen according to the sizes of a standard word embedding values, which was obtained from training with large text corpus, for both English and Portuguese languages, and they are publicly available.

3) Word dictionary: At this step, a word dictionary is created, assigning an identification number to each word. Here, the text of each news is converted into a list of integers, representing the identifier of each word. After this step, all news' text is transformed into a one-size-fits-all string, according to the average of words of the entire dataset. This way, texts with the number of words smaller than the defined length will be filled with 0's. This step of transforming all texts to a single size is necessary to train the model in batches.

4) Dataset division: Finally, we use the holdout method to divide the dataset into two subsets at random. The first one is used just for training (to adjust the LSTM model), whilst the second one is used for testing the model classification performance. The split ratio is 80% for training, and 20% for testing.

B. Model architecture

During this step, the model architecture is built. The model is an Artificial Neural Network composed of three layers: an input layer (embedding) that has the task of loading the pre-trained vector of each news word, a LSTM layer with output dimensionality equal to the dimensionality size of the input vector (50, 100 and 300), and a dense layer with output size 2, with the softmax activation function. The loss function chosen is the binary cross-entropy due to that the classification is binary, and the chosen optimiser is RMSProp.

C. Training and Test

At this stage, the samples from the training subset are crossed through the model for n training epochs, until the model reaches the appropriate level of convergence. After- wards, the model is tested with samples from the test subset only.

D. Evaluation

To evaluate the proposed model, we use the confusion matrix, whose indicators are: • TP: the prediction is as fake news and text is about fake news. • FP: the news is predicted as fake news, but the text is about a true news. • TN: the prediction is as true news and the text is true news too. • FN: the news is predicted as true news, but the text is fake news. The metrics are computed using the confusion matrix indicators (precision, recall, f1-score and accuracy) . These are used to evaluate the performance of the model. Besides, the mentioned metrics are be used to evaluate the different word embedding methods for the classification, as well as to compare their results with those from other works.

Dataset

The experiments that were performed in this work were conducted using the Fake.Br dataset. This dataset contains 7,200 news, with 3,600 being fake news and 3,600 being true news. The news were analysed and extracted from news websites, with the publication period being between January 2016 and January 2018. The set of news composed of fake news from the Fake.Br corpus was collected manually, while the news that form part of the true news sub-dataset were collected in a semi automatic manner. The first stage of the collection was performed using a web crawler8, seeking news based on keywords from a fake news story. After this step, for each fake news, a measure of lexical similarity with the real news collected was applied, choosing the most similar to the fake news. The last step of the collection of the true news was a manual check to ensure that for every true news there would be a fake one related to the same subject. The dataset can be broken down into 6 categories related to their main subjects: politics, TV and celebrities, society and daily news, science and technology, economics and religion.

Results and discussion

The objective of this study is to verify the classification performance of the model using the representation generated through fastText, Glove, and word2vec; as well as verifying the performance of the architecture using LSTM. All classification performance metrics used in this work are computed from the confusion matrix. Model training was done with the 7,200 news dataset, with a split of 80% for training and 20% for testing. The dataset split process was performed 10 times, so each method was performed 10 times with different datasets and testing.

Conclusion

Due to the growing concern with impact that fake news may cause to the society, it emerges the need for automated methods for detecting fake news using computational approaches. Methods such as DL and NLP have been widely used to assign categories to texts according their content. In this sense, distributed representations of words and RNNs are widely used for text classification tasks. Thus, in order to contribute to the detection of fake news in texts, this work focused on obtaining distributed representation from words with three methods of word embedding, and thus classifying news using type-LSTM RNNs. At the experiment conducted to verify the classification performance concern word embedding methods, it was observed that these methods are capable of capturing context- based characteristics. The best classification performance was obtained using the vectors generated by the Word2vec method, however, the results obtained with the other studied methods were similar. At this sense, results suggest that the proposed approach is promising and may be used as a tool for ranking fake news. Regarding the results obtained from conducted experiment using embedding layer with a dimensionality of 300, these were better when compared with the dimensions 50 and 100, just for the Word2vec. In comparison of our architecture with different others pro- posed in literature it is possible to verify similar results, despite others are more deep (and more sophisticated) architectures, obtaining good classification performance with average value about 0.9515 to the f1-score metric. The architecture proposed at this work seems to be better in classification performance when compared to simpler architectures. For training and test step, a publicly available dataset containing true and fake news in Portuguese was used. From obtained results it is possible to evaluate that the suggested approach using word embedding methods together with LSTM is capable of classifying fake news at Portuguese language story. Overall, from results, we believe that our approach is very promising for detecting fake news on news stories, and it could be generalised to other news corpus.

As future works, a possible research direction is to apply new methods for processing natural language, such as encoder architectures based on Transformer DL model such as the Bidirectional Encoder Representations from Transformers (BERT). BERT is a model already trained in a large text base and which can be retrained on new data (like as the problem in question) in order to learn the difference between the general meaning of a word and the meaning of the word for a specific context. This issue is specially important for fake news classification, because the context understanding is a not trivial task and the main objective for a good classification of fake news is to minimise the semantic gap between general meaning and specific meaning of words. Another possible direction of research is the detection of fake news from text on images, thus using a optical character recognition early.

NATURAL LANGUAGE PROCESSING

Abstract

Fake news detection is a critical yet challenging problem in Natural Language Processing (NLP). The rapid rise of social networking platforms has not only yielded a vast increase in information accessibility but has also accelerated the spread of fake news. Thus, the effect of fake news has been growing, sometimes extending to the offline world and threatening public safety. Given the massive amount of Web content, automatic fake news detection is a practical NLP problem useful to all online content providers, in order to reduce the human time and effort to detect and prevent the spread of fake news. In this paper, we describe the challenges involved in fake news detection and also describe related tasks. We systematically review and compare the task formulations, datasets and NLP solutions that have been developed for this task, and also discuss the potentials and limitations of them. Based on our insights, we outline promising research directions, including more fine grained, detailed, fair, and practical detection models. We also highlight the difference between fake news detection and other related tasks, and the importance of NLP solutions for fake news detection.

Introduction

Automated fake news detection is the task of assessing the truthfulness of claims in news. This is a new but critical NLP problem because both traditional news media and social media have huge social-political impacts on every individual in the society. For example, expo- sure to fake news can cause attitudes of inefficeancy, alienation, and cynicism toward certain political candidates (Balmas, 2014). Fake news even relates to real-world violent events that threaten public safety (e.g., the Pizza Gate (Kang and Goldman, 2016)). Detecting fake news is an important application in the world that NLP can help with, as it also creates broader impacts on how technologies can facilitate the verification of the veracity of claims while educating the general public. The conventional solution to this task is to ask professionals such as journalists to check claims against evidence based on previously spoken or written facts. However, it is time- consuming and expensive. For example, PolitiFact1 takes three editors to judge whether a piece of news is real or not. As the Internet community and the speed of the spread of information are growing rapidly, automated fake news detection on internet content has gained interest in the Artificial Intelligence research community. The goal of automatic fake news detection is to reduce the human time and effort to detect fake news and help us stop spreading it. The task of fake news detection has been studied from various perspectives with the development in subareas of Computer Science, such as Machine Learning (ML), Data Mining (DM), and NLP. In this paper, we survey automated fake news detection from the perspective of NLP. Broadly speaking, we intro- duce the technical challenges in fake news detection and how researchers define different tasks and formulate ML solutions to tackle this problem. We discuss the pros and cons, as well as the potential pitfalls and drawbacks of each ask. More specifically, we provide an overview of research efforts for fake news detection and a systematic compari- son of their task definitions, datasets, model construction, and performances. We also discuss a guideline for future research in this direction. This paper also includes some other aspects such as social engagement analysis. Our contributions are three-fold:

- We provide the first comprehensive review of Natural Language Processing solutions for automatic fake news detection;

- We systematically analyze how fake news detection is aligned with existing NLP tasks, and discuss the assumptions and notable issues for different formulations of the problem;

- We categorize and summarize available datasets, NLP approaches, and results, providing first hand experiences and accessible introductions for new researchers interested in this problem.

LIAR

LSTM based models achieve higher accuracy than CNN based models. The additional meta-data is also important. Karimi et al. (2018) supplement LIAR by adding the verdict reports written by annotators and raise accuracy by 4%. Kirilin and Strube (2018) improve accuracy by 21% through replacing the credibility history in LIAR with a larger credibility source (speaker2credit7). The two papers also show the attention scores for verdictre ports/speaker credit are higher than the statement of claim.

Datasets

A significant challenge for automated fake news detection is the availability and quality of the datasets. We categorize public fake-news datasets into three categories: claims, en- tire articles, and Social Networking Services (SNS) data. Claims are one or a few sentences including information worth validating , while en- tire articles are composed of many sentences related to each other constituting information as the whole. SNS data are similar to claims in length but featured by structured data of accounts and posts, including a lot of non-text data.

1) Claims POLITIFACT, CHANNEL4.COM2, and SNOPES3 are three sources for manually labeled short claims in news, which is collected and labeled manually. Editors handpicked the claims from a variety of occasions such as debate, campaign, Facebook, Twitter, interviews, ads, etc. Many datasets are created based on these websites. Vlachos and Riedel released the first public fake news detection dataset gathering data from POLITIFACT and CHANNEL4.COM. This dataset has 221 statements with the date it was made, the speaker and the URL, and the veracity label of a five point scale. EMERGENT (Ferreira and Vlachos) is the early work of claim- verification dataset too. It is for stance classification in the context of fact-checking, including claim with some documents for or against them. This dataset can improve fact-checking in the condition that some articles related to the claim were given. Vlachos includes only 221 claims and Emergent includes only 300 claims so that it was impractical to use them for machine learning based assessments. These days, datasets with many claims are published, which can use as an improved version of the first two. A recent benchmark dataset for fake news detection is LIAR (Wang, 2017). This dataset collected data from Politi fact as Vlachos and Riedel (2014), but includes 12,836 real world short statements, and each statement is labeled with six-grade truthfulness. The information about the subjects, party, context, and speakers are also included in this dataset. For the datasets from Politifact articles, Rashkin et al. (2017) also published large datasets. They collect articles from (Politifact's spin-off site) too. Fever (Thorne et al., 2018) is a dataset providing related evidence for fact checking. In this point, it is similar to EMERGENT. Fever contains185,445claims generated from Wikipedia data. Each statement is labeled as Supported, Refuted, or Not Enough Info. They also marked which sentences from Wikipedia they use as evidence. Fever makes it possible to develop a system that can predict the truthfulness of a claim together with the evidence, even though the type of facts and evidence from Wikipedia may still exhibit some major stylistic differences from those in real-world political campaigns.

2) Entire-Article Datasets There are several datasets for fake news detection predicting whether the entire article is true or fake. For example, FAKENEWSNET (Shu et al., 2017a; Shu et al., 2017b; Shu et al., 2018) is an ongoing data collection project for fake news research. It consists of headlines and body texts of fake news articles based on BuzzFeed and PolitiFact. It also collects information about the social engagements of these articles from Twitter. BS DETECTOR4 is collected from a browser extension named BS Detector, indicating that its labels are the out- puts of the BS Detector, not human annotators. BS Detector searches all links on a web page at issue for references to unreliable sources by checking against a manually com- piled list of unreliable domains.

3) Posts On Social Networking Services There are some datasets for fake news detection focusing on

SNS, but they tend to have a limited set of topics and can be less related to news. BUZZFEEDNEWS5 collects 2,282 posts from 9 news agencies on Facebook. Each post is factchecked by 5 BuzzFeed journalists. The advantages of this dataset are that the articles are collected from both sides of left leaning and right-leaning organizations. There are two enriched versions of BUZZFEEDNEWS: Potthast et al. (2017) enriched them by adding data such as the linked articles, and BUZ- ZFACE (Santia and Williams, 2018) extends the BuzzFeed dataset with the 1.6 million comments related to news articles on Facebook. SOME-LIKE-IT-HOAX6 (Tacchini et al., 2017) consists of 15,500 posts from 32 Facebook pages, that is, the public profile of organizations (14 conspiracy and 18 scientific organizations). This dataset is labeled based on the identity of the publisher instead of post-level annotations. A potential pitfall of such a dataset is that such kind of labeling strategies can result in machine learning models learning characteristics of each publisher, rather than that of the fake news. PHEME (Zubiaga et al., 2016) and CREDBANK (Mitra and Gilbert, 2015) are two Twitter datasets. PHEME contains 330 twitter threads (a series of connected Tweets from one person) of nine newsworthy events, labeled as true or false. CREDBANK contains 60 million tweets covering 96 days, grouped into 1,049 events with a 30-dimensional vector of truthfulness labels. Each event was rated on a 5-point Likert scale of truthfulness by 30 human annotators. They concatenate 30 ratings asa vector because they find it difficult to reduce it to a one-dimensional score. As mentioned above, these datasets were created for verifying the truthfulnessof tweets. Thus they are limited to a few topics and can include tweets with no relationship to news. Hence both datasets are not ideal for fake news detection, and they are more frequently used for rumor detection

Regression

Fake news detection can also be formulated as a regression task, where the output is a numeric score of truthfulness. Usually, evaluation is done by calculating the difference between the predicted scores and the ground truth scores Pearson/Spearman Correlations.

However, since the available datasets have discrete ground truth scores, the challenge here is how to convert the discrete labels to numeric scores.

Rhetorical Approach

Rhetorical Structure Theory (RST), sometimes combined with the Vector Space Model (VSM), is also used for fake news detection (Rubin et al., 2015b; Della Vedova et al., 2018; Shu et al., 2017b). RST is an analytic framework for the coherence of a story. Through defining the semantic role (e.g., a sentence for Circumstance, Evidence, and Purpose) of text units, this framework can systematically identify the essential idea and analyze the characteristics of the input text. Fake news is then identified according to its coherence and structure. To explain the results by RST, VSM is used to convert news texts into vectors, which are compared to the center of true news and fake news in high-dimensional RST space. Each dimension of the vector space indicates the number of rhetorical relations in the news text.

Neural Network Models

Recurrent Neural Network (RNN) is very popular in Natural Language Processing, especially Long Short Term Memory (LSTM), which solves the vanishing gradient problem so that it can capture longer-term dependencies. In Section 6., many models based on LSTM per- form high accuracy on LIAR and FEVER. In addition, Rashkin et al. (2017) set up two LSTM models and input text as simple word embeddings to one side and as LIWC feature vectors to the other. In both cases, they were more accurate than NBC and Maximum Entropy(MaxEnt) models, though only slightly. Convolutional neural networks (CNN) are also widely used since they succeed in many text classification tasks. Wang (2017) uses a model based on Kim's CNN (Kim, 2014), concatenating the max-pooled text representations with the meta-data representation from the bi-directional LSTM. CNN is also used for extracting features with a variety of meta-data. For example, Deligiannis et al. (2018) took graph-like data of relation- ships between news and publishers as input for CNN and assess news with them.

Non-Neural Network Models

Support Vector Machine (SVM) and Naive Bayes Classifier (NBC) are frequently used classification mod els (Conroy et al., 2015; Khurana and Intelligentie, 2017; Shu et al., 2018).These two models differ a lot in structure and both of them are usually used as baseline models. Logistic regression (LR) (Khurana and Intelligent, 2017; Bhattacharjee et al., 2017) and decision tree such as Random Forest Classifier (RFC) (Hassan et al., 2017) are also used occasionally.

Machine Learning Models

As mentioned in Section 3., the majority of existing re- search uses supervised methods while semi supervised or unsupervised methods are less commonly used. In this section, we mainly describe classification models with several actual examples.

Conclusion

In this survey, we first reveal the importance and definitions of automatic fake news detection. Then we compare and discuss the most recent benchmark datasets and experimental results of different methods. Based on our observations, we propose new recommendations for future datasets, and also give the following suggestions for our future fake news detection model: investigate whether the hand crafted features can be combined with neural network models, appropriate usage of non-textual data, and extending the way of verification with contents.