LOG BOOK

October 22, 2025 2:21pm

Started: October 22,2025 2:21pm

Research Question/ Problem:

What is misinformation?

Jan 3, 5:18pm

In our world today anyone can get misinformed by people, social media, newspapers, books, school or from yourself. Misinformation is recognized as fake news, leading the public into the wrong direction of something which is not true causing hate speech and reducing the trust in the democratic processes. Although the plus side about misinformation is that it's not done with an intention of causing harm, however it does not prevent the confusion around people and what they know, which puts them in a position of who's right or wrong with not an exact answer. This type of confusion leads people with not the right sourced answers or information misleading others into thinking it’s right. For example your friend informs you that chocolate is really unhealthy and should never have it because they probably read something online which was not true. The reason behind this is because your friend didn’t mean to harm you intentionally, but for you to be safe. Although, the friend is wrong because chocolate is not bad completely unless you consume it more than you should or else chocolate can be good for your health. It’s very important to have the right knowledge and information so you don’t mislead into thinking something else. Also being cautious of where you get your information is crucial.

Nowadays younger kids are vulnerable and naive to propaganda and misinformation because they were born with lots of technology around them, and they spend most of their time watching their phone, television, chatting online, as well as playing games. Since they are actively on technology they learn many things which are false and true which can mislead into thinking it’s true. AI has developed a lot of ways to trick your mind and fool people easily, but it can also be used for good.

Interest towards topic: March 3, 2026 4:36pm
The reason why I was directed into this topic is because misinformation has been guiding people in the wrong direction, and people are being misinformed about what they watch, read, listen, and share. Unintentionally people get misguided and the reason could be because of social media, fake news and websites. To tackle this I chose computer science because it can identify misinformation and will help the person by identifying real or fake news. Computer science can help solve real world problems and I decided to create a step-by-step guide to detect misinformation so these types of problems can be tackled.
Additionally, I chose python as an effective non time consuming language for a computer program because it takes less time to recognize and process information and provides you with credible sources of information without getting misinformed.

Why is using python to create a computer program helpful to detect misinformation? March 3, 2026 5:36pm

Creating a computer program incorporating python to detect misinformation is efficient because it can process many articles or news stories at once in less time. Python is a powerful language which is commonly used for data analysis workflows. Python is popular in the community of machine learning, which provides many accessible libraries for machine learning algorithms as well as other tasks such as splitting datasets, training and building a program.
A computer program using python can be used to recognize a specific text from another. In the past few years scientists have become active in creating algorithms to detect misinformation, this way it helps people to understand fake news.
Computers are able to examine writing styles from ones which are in a deceptive text and ones which are written honestly. Advancement in machine learning has made it possible for computers to complete tasks which would take humans a longer amount of time. Also artificial intelligence makes it easy to train algorithms to detect and provide credible information from real articles.

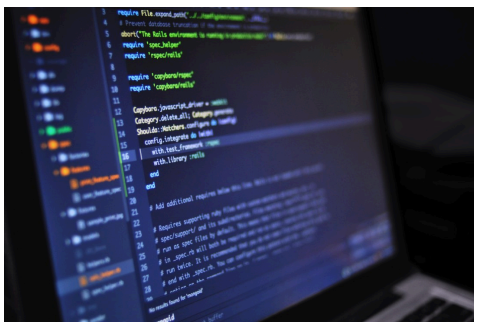
Is computer science the only solution to detect misinformation? March 3, 2026 5:36pm

No, Computer science is not the only way to detect misinformation. There are techniques and skilled ways to handle problems regarding fake news, such as human approaches,psychological strategies, and literacy behavior towards information. Critically thinking and paying attention towards how different types of information is presented helps to separate the false and true info. Experience towards literacy can help one to investigate real and fake information on articles or websites. Critical thinking is essential for recognizing non-credible sources. Another method to identify misinformation is by prebunking or pre-emptive debunking, seeking to help people from being misinformed in the first place. The common method for prebunking is psychological inoculation, which means exposure to weak versions of falsehood forms resistance to future persuasion. Prebunking can reduce susceptibility of falsehoods and larger techniques incorporated to manipulate information. Current evidence also suggests that prebunking is practical in real world settings such as social media.

How is Computer science helpful?

Jan 6, 7:06pm

Computer science is helpful because it's the study of computers and how its abilities can solve real life problems, like helping surgeons perform surgeries. It helped to navigate a spacecraft to the moon and create navigational systems so it's easier knowing where to go following directions which helped to solve our problems and make our lives easier to cope with. Computer science adds effects to videos, movies, as well as games so us humans can have entertainment. Computer scientists use and create technology to solve problems, they program devices or computers to finish tasks much faster and efficiently and you can teach them new things. This means becoming a computer scientist is to switch from being a consumer to a creator. Computers are powerful but not that smart and are able to do what they have been told in the form of instructions or algorithms for a computer to understand. An algorithm is a list of step by step instructions of what the program should do. Computer science is the process of solving technical problems as well as complex organizational problems. Computer science also provides a vast number of jobs for people to choose from, forming many opportunities and access to becoming a software engineer, systems administrator, data scientist, computer scientist, web developer, data analyst, and more. Without these jobs or interests it would be difficult to have access to games, movies,websites, platforms, or apps made by coding and algorithms for the technology to understand. Computer science is helpful because it teaches us how to code, make games, solve world wide problems efficiently and in the right way and we get to learn how your favourite apps are made, games which you consume, and movies which you watch, and music you listen to will teach you how the computer functions and what language it understands just like humans. Computer science is helpful because if you're in a situation where you get a ransomware attack it's good to know some computer science skills in order to tackle these types of difficulties without not being able to do anything and overcome them by using coding. Honestly I think that computer science is a very good skill to learn to solve problems and to tackle ransomware, spams or hacks, and mainly using this skill for the right reason and right thing, not to harm anyone or take advantage of the knowledge.



Data/ Steps to creating a computer program using python: Feb 25, 12:55pm

1. To begin with you should decide on a computer language which will effectively guide you to detect misinformation.
A suitable language you can use is Python. Python is rooted with many advantages. Python is well-structured for the environment of AI development and expansive tech stack of libraries and frameworks. It remarkably reduces SDLC(software development life cycle). SDLC is a process for training, building, testing, and maintaining software. Instead of using SDLC Python makes it manageable for developers so they don't have to code from the beginning. Several Python libraries are suitable for the AI-ML development bill, as they are able to access, cooperate,and transform difficult data in large volumes.

These libraries include:

NumPy: performs mathematical and statistical operations

SciPy: Used for scientific and technical computing

Theano: Building deep learning models

Pandas: For high-level data analysis

PyBrain: Used for neutral networks

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
import re
import string
```

2. **Choose a Dataset:**
For building a classifier, we would first need to have a good dataset. In order for a data set to have good categories it would have to be ‘fake’ and ‘real’ news. The next step would be to have a big enough sample size to avoid overfitting and to be split into testing datasets.

Split data:

```
from sklearn.model_selection import train_test_split
x_train ,x_test ,y_train ,y_test= train_test_split(x,y,test_size=0.30)
```

You can use some real news articles, some which are fake. In order to do that you can use a dataset from kaggle uploaded by Emine Bozkus. This dataset has over 20,000 articles which are fake and real news.

3. **Detecting fake news with Python**
First you would have to download the datasets from above. You can name real news real.csv and fake news fake.csv. CSV files are compatible for storing and sharing data. Then you’ll use libraries which will be used in the steps ahead.

Pandas: For high-level data analysis

Nltk: Natural Language Tool Kit; preprocessing news article texts

Scikit_learn: To build and evaluate fake news detection model

```
df_fake=pd.read_csv('Fake.csv')
df_true=pd.read_csv('True.csv')
```
\\ Reading the dataset.

```
df_fake_mantest=df_fake.tail(10)
for i in range(23480,23470,-1):
    df_fake.drop([i],axis=0,inplace=True)
df_true_mantest=df_true.tail(10)
for i in range (21416,21406,-1):
    df_true.drop([i],axis=0,inplace=True)
```
\\ Removing last 10 entries.

```
df_mantest= pd.concat([df_fake_mantest,df_true_mantest],axis=0)
df_mantest.to_csv("mantest.csv")
```
\\ CSV file for manual testing.

```
df_merge=pd.concat([df_fake,df_true],axis=0)
df_merge.head(10)
```
\\ CSV file for Model Predictions

4. **Data processing for Fake news implementing Python**
First you can read the real news real.csv and fake news fake.csv. For this part you can use Pandas

Normalizing the text:
Next, preprocess the text in the news articles so it becomes easier for the model to analyze. Mainly, you'll be removing some texts which don't help in classifying it as real or fake. Now you will do some common string processing to simplify the text. The basic string process is a sequence of characters. It's a method to examine individual characters such as extracting parts of strings and cleaning unnecessary data. After that's done, you can use the nltk library to do some processing tasks.

```
def cleaning(text):
    text=text.lower()
    text = re.sub('\[.*?\]', '', text)
    text = re.sub("\\W"," ",text)
    text = re.sub('https?://\S+|www\.\S+', '', text)
    text = re.sub('<.*?>+', '', text)
    text = re.sub('[%s]' % re.escape(string.punctuation), '', text)
    text = re.sub('\n', '', text)
    text = re.sub('\w*\d\w*', '', text)
    return text
```

In the code you then apply an article_preprocessor function to the entire dataset. In the function, you start by converting the text into lowercase letters and then remove digits as well as punctuation. Then use the nltk library to simplify the text more using NLP techniques.

1. Tokenize the text into a list of words
2. Filter the list by removing stop words such as ‘the’, ‘in’, ‘as’, ‘for’ etc.
3. Stem the words to shorten them to their root by taking out affixes. Such as, ‘intelligence’ and ‘intelligent’ would become ‘intellig’. This way it makes it easier for the algorithm to understand and process.

   Feature Engineering
   Now, you have datasets which are cleaned and normalized. However, the machine learning algorithms that you’ll be using don’t fully understand strings; they understand vector data. So, you’ll be turning the text data into vectors, which can be used as inputs to transform, test, and hypothesize machine learning algorithms.

For the algorithm, you’ll be converting every article text into a vector where each word token appearing in the article turns into a dimension and the token count becomes the corresponding magnitude. This could be referred to as the bag-of-words approach which is very simple. The sklearn provides a CountVectorizer class to accomplish this. In the code you split the dataset into testing and training sets by incorporating sklearn . Model_selection . train_test_split function. You pass the text and label columns as inputs and outputs. Also you specify a test_size=0.2, which refers to a random 20% of data will be used for testing, as the rest will be used for training. To make sure the randomly selected testing and training splits stay consistent between successive runs, you specify a random_state.

After splitting the dataset, you acquire the training and testing inputs x_train and x_test and expected outputs y_train and y_test. Then you vectorize the inputs using the CountVectorizer.

5. **Training the fake news detection model**
From the last step, you would now have a dataset of vectorized news articles and corresponding labels (‘real’ or ‘fake’). Now that you use this, you will be building a predictive model that can classify an article as ‘real’ or ‘fake’. For demonstration, you can use the naive bayes algorithm.
In the code you have created, you used sklearn . naive_bayes .MultinomialNB class to begin a classifier, then you fitted the training data in the classifier. By doing this it completes the training, and the model is ready to be in use.

```
def output_lable(n):
    if n==0:
        return "Fake News"
    elif n==1:
        return "Not a Fake news"
def manual_testing(news):
    testing_news={"text":[news]}
    new_def_test=pd.DataFrame(testing_news)
    new_def_test["text"]=new_def_test["text"].apply(cleaning)
    new_x_test=new_def_test["text"]
    new_v_test=vec.transform(new_x_test)
    lpred=lr.predict(new_v_test)
    dpred=dt.predict(new_v_test)
    gpred=gbc.predict(new_v_test)
    rpred=rfc.predict(new_v_test)
```

```
news=str(input())
manual_testing(news)



LR Prediction: Fake News
DT Prediction: Fake News
GBC Prediction: Fake News
RFC Prediction: Fake News
```

6. **Evaluating the model**
Lastly you will be running the model on test inputs and comparing the hypothesized outputs with the real outputs. You'll be measuring the performance using metrics accuracy, precision , recall score, and F1 score. These can be calculated using methods of the sklearn library.

Method: March 3, 2026 7:09pm

1. Questions which were asked to google:

What is misinformation?

Why is using python to create a computer program helpful to detect misinformation?

Is computer science the only solution to detect misinformation?

How is Computer science helpful?

2. Citations for each of these questions:

What is misinformation? How is computer science helpful?
https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propaganda-misinformation

https://www.apa.org/topics/journalism-facts/how-why-misinformation-spreads

https://csuglobal.edu/blog/why-computer-science-so-important

https://www.cmich.edu/blog/all-things-higher-ed/23-careers-computer-science-degree

https://pirg.org/edfund/articles/misinformation-on-social-media/

Why is using python to create a computer program helpful to detect misinformation?
https://www.newscatcherapi.com/blog-posts/fake-news-detection-using-python#:~:text=Hence%2C%20writing%20a%20computer%20program,using%20Python%20for%20this%20tutorial.

https://theconversation.com/the-language-gives-it-away-how-an-algorithm-can-help-us-detect-fake-news-120199#:~:text=For%20example%2C%20there%20are%20computer,compared%20to%20one%20written%20honestly.

https://www.pbs.org/newshour/science/the-language-gives-it-away-how-an-algorithm-can-help-us-detect-fake-news#:~:text=Algorithms%20can%20help%20detect%20misinformation%20by%20using,*%20**Identifying%20linguistic%20characteristics%20of%20fake%20news**

Is computer science the only solution to detect misinformation?
https://pmc.ncbi.nlm.nih.gov/articles/PMC7134234/#:~:text=Abstract,evaluating%20the%20plausibility%20of%20arguments.

https://www.apa.org/topics/journalism-facts/misinformation-interventions#:~:text=research%20is%20required.-,Prebunking,to%20forestall%20misinformation%20at%20scale.

3. Research related to making a guide using websites to show the steps:

After the research was complete I began working on the data and the steps to making a computer program using python, as well as explained how to build a program properly and what should be done in order to function and effectively work.

4. The last step: Conclusion

After completing research and data I then finished my conclusion and provided information of what went well and what I could've done in order to make a computer program physically.

Conclusion:

In conclusion this step-by-step guide will help people to make their own computer programs to verify the information they collect from various sources. In addition it will give a chance for people to learn about Python and its different libraries to help it function, as well as being able to explore computer science and its multiple languages. All in all I was glad that I managed to find sources which effectively helped me create a guide to detect true and false information in a simple manner and made it easier to understand how to create your own computer program. However, if I were to try something different with this research and method I would try to create a computer program and learn how to test and build an effective model.

Citations For Research:

What is misinformation? How is computer science helpful?

https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propaganda-misinformation

https://www.apa.org/topics/journalism-facts/how-why-misinformation-spreads

https://csuglobal.edu/blog/why-computer-science-so-important

https://www.cmich.edu/blog/all-things-higher-ed/23-careers-computer-science-degree

https://pirg.org/edfund/articles/misinformation-on-social-media/

A computer science book was used in order to answer questions for How is computer science helpful?

Why is using python to create a computer program helpful to detect misinformation?

https://www.newscatcherapi.com/blog-posts/fake-news-detection-using-python#:~:text=Hence%2C%20writing%20a%20computer%20program,using%20Python%20for%20this%20tutorial.

https://theconversation.com/the-language-gives-it-away-how-an-algorithm-can-help-us-detect-fake-news-120199#:~:text=For%20example%2C%20there%20are%20computer,compared%20to%20one%20written%20honestly.

https://www.pbs.org/newshour/science/the-language-gives-it-away-how-an-algorithm-can-help-us-detect-fake-news#:~:text=Algorithms%20can%20help%20detect%20misinformation%20by%20using,*%20**Identifying%20linguistic%20characteristics%20of%20fake%20news**

Is computer science the only solution to detect misinformation?
https://pmc.ncbi.nlm.nih.gov/articles/PMC7134234/#:~:text=Abstract,evaluating%20the%20plausibility%20of%20arguments.

https://www.apa.org/topics/journalism-facts/misinformation-interventions#:~:text=research%20is%20required.-,Prebunking,to%20forestall%20misinformation%20at%20scale.

Citations For Data:

https://www.newscatcherapi.com/blog-posts/fake-news-detection-using-python#:~:text=For%20building%20any%20classifier%2C%20we,or%20a%20different%20dataset%20altogether.

https://medium.com/gatorhut/identifying-fake-news-using-real-time-analytics-in-python-43cacf94c4f3

https://dev.to/piushopkins/detecting-fake-news-with-python-and-machine-learning-4mh5#:~:text=This%20Python%20project%20module%20for,Lab%20to%20run%20your%20codes.

https://www.damcogroup.com/blogs/why-experts-prefer-python-for-ai-ml-development

https://www.atlassian.com/agile/software-development/sdlc

https://www.adobe.com/acrobat/resources/document-files/what-is-a-csv-file.html

https://processing.org/reference/String.html

Images:

https://futurumcareers.com/misinformation-what-it-is-and-how-to-spot-it

https://www.security.org/digital-security/misinformation-disinformation-survey/

https://www.linkedin.com/pulse/top-10-reasons-why-you-should-study-computer-science-kisan-tamang

https://medium.com/gatorhut/identifying-fake-news-using-real-time-analytics-in-python-43cacf94c4f3

With the help of AI overview I was able to find websites which I can source and use the information for my project.

The images for the data I sourced, shows after doing each step what would it look like as its result.