Problem Statement

Infinity Influencer Marketing

is a brand new online marketing firm and they want us to build a Java/Python microservice to identify whether a given post_text is sponsored or not.

- post_text is a english language text which may contain emoji (e.g 😀) or special characters
- To determine whether a post_text is sponsored or not we would like to use the following logic:
    - If below keywords are present in the post_text, we say that post_text is a sponsored post
        - #ad
        - #sponsored
        - advertisement
    - Non-sponsored otherwise
- We should be able to add new keywords to the original set of keywords (#ad, #sponsored etc) so that our vocabulary is not static
- There are no restrictions on usage of any open source frameworks e.g Flask, Spring, Hibernate etc
- There is no specific requirements to build any form of user interface for this assignment
- The candidate can make appropriate assumptions as required. Please mention them in the submission

Bonus Points

- Bonus points are awarded if
    - We use some kind of database (in-memory store) to persist the **vocabulary**
    - Create ORM utility to Insert and Select records from the table

Web services

The microservice should host the following endpoints:

HTTP Method	URL Pattern	Function
GET	/api/vocab	Return list of keywords (e.g #ad, #sponsored, advertisement) used to determine whether or not post_text is sponsored
POST	/api/vocab	Add new keyword in already existing vocab
POST	/api/prediction	Predicts if the given post_text is sponsored or non-sponsored by checking for vocab keywords in the given post_text. If any of the vocab keywords is present, the service should predict “sponsored”. If none of the vocab keywords are present, the service should predict“non-sponsored”

- Note: **POST** call should use **content-type** header value as **application/json**

Request and Response Format

- /api/vocab

Accepts `GET request`

Request format

```
/api/vocab
```

Response format

```
{
    "vocab": [
        "#ad",
        "#sponsored",
        "advertisement"
    ]
}
```

- /api/vocab

Accepts `POST request`

Request format

```
{
    "vocab": [
        "ambassador"
    ]
}
```

Response format

```
{
    "vocab": [
        "#ad",
        "#sponsored",
        "advertisement",
        "ambassador"
    ]
}
```

- /api/prediction

Accepts `POST request`

Request format 1

```
{
    "post_text": "#ad Love these cool toys at #ToysRUs. Go check them out"
}
```

Response format 1

```
{
    "prediction": "sponsored"
}
```

Request format 2

```
{
    "post_text": "My new year resolution is to stay fit and healthy"
}
```

Response format 2

```
{
    "prediction": "non-sponsored"
}
```

Deliverable

GitHub Deliverable

- To allow for the easy evaluation of your work, please submit a github link to the repository. (If one does not have a GITHUB account, then a zipped version of the project folder is acceptable in that case)
- It should contain either of the following files
    - requirements.txt if PYTHON
    - pom.xml if JAVA maven project
    - Or any other file with all dependencies listed so that we can build the project successfully
- Readme.md file should explain how to run the code and call the endpoints
- There is no specific requirement to build any user interface

Bonus Points

- Bonus points are awarded if
    - Candidate provides a dockerfile/docker image url which can be used to run the code