

## Task1

### JSON File Uses

JSON (JavaScript Object Notation) is a lightweight data interchange format that is easy for humans to read and write and easy for machines to parse and generate. Here are some common uses of JSON files:

1. Data exchange format: JSON is often used as a format for exchanging data between web servers and clients. JSON files can be used to store and transmit structured data such as user preferences, product information, and financial data.
2. Configuration files: JSON files can be used to store configuration settings for web applications and software. This allows developers to easily configure and deploy applications across multiple environments.
3. API responses: Many web APIs return JSON responses to requests. This is because JSON is a lightweight format that can be easily parsed by a variety of programming languages.
4. Logging: JSON files can be used to store logs for applications and systems. This makes it easy to search and analyze log data using tools like Elasticsearch and Kibana.
5. Testing: JSON files can be used for testing applications and APIs. Test data can be easily stored in JSON files and used to test application functionality.
6. Serialization: JSON files can be used for object serialization, which is the process of converting objects in memory to a format that can be stored or transmitted. This is useful for persisting application data or sending objects over the network.

Overall, JSON files are versatile and widely used in modern software development.

## Task2

### what are the online services that provide us access on GPUs

There are several online services that provide access to GPUs, which can be useful for running machine learning and deep learning workloads. Here are some popular ones:

1. Amazon Web Services (AWS): AWS provides a variety of GPU instances, including instances with NVIDIA Tesla GPUs optimized for deep learning workloads. These instances can be launched through the AWS console or via the command line using tools like the AWS CLI.
2. Google Cloud Platform (GCP): GCP also provides a variety of GPU instances, including

instances with NVIDIA GPUs optimized for deep learning workloads. These instances can be launched through the GCP console or via the command line using tools like the gcloud CLI.

3. Microsoft Azure: Azure provides GPU instances with NVIDIA GPUs optimized for deep learning workloads. These instances can be launched through the Azure portal or via the command line using tools like the Azure CLI.

4. IBM Cloud: IBM Cloud provides GPU instances with NVIDIA GPUs optimized for deep learning workloads. These instances can be launched through the IBM Cloud console or via the command line using tools like the IBM Cloud CLI.

5. Paperspace: Paperspace is a cloud computing platform that provides access to powerful GPU instances for machine learning and other workloads. Users can launch GPU instances via a web-based console or via the command line.

These online services provide users with on-demand access to powerful GPU resources, which can be especially useful for running computationally intensive workloads. Each service has its own pricing and usage policies, so it's important to research and compare different options to find the one that best fits your needs and budget.

6. Colab: Colab is a free online service provided by Google that allows users to run Jupyter notebooks with access to GPUs and TPUs (tensor processing units). It's a popular platform for machine learning education and research, and it integrates with other Google services like Drive and GitHub.

7. FloydHub: FloydHub is a cloud platform for deep learning that provides GPU instances for running workloads. It also includes tools for managing data and running experiments, and it integrates with popular deep learning frameworks like TensorFlow and PyTorch.

8. Crestle: Crestle is a cloud platform that provides access to GPUs and other resources for running machine learning workloads. It includes pre-installed software packages and tools for data processing and visualization.

9. Paperspace Gradient: Paperspace Gradient is a cloud platform that provides access to GPUs and TPUs for running machine learning workloads. It also includes tools for managing data and running experiments, and it integrates with popular deep learning frameworks like TensorFlow and PyTorch.

10. Alibaba Cloud: Alibaba Cloud provides GPU instances for running machine learning and other workloads. It includes a variety of GPU options, including NVIDIA Tesla and AMD Radeon GPUs, and it integrates with Alibaba Cloud's other services for storage, networking, and security.

These are just a few examples of the many online services that provide access to GPUs. Depending on your specific needs and budget, there may be other services that are better suited for your use case.

### **Task3**

#### **Extentions that block websits from tracking users**

- Privacy Badger
- UltraBlock
- Incogation Adblocker
- Ghostery\_Orivacy Ad Blocker
- DuckDuckGo Privacy Essentials
- Disconnect

### **Task4**

#### **What is Data Mining?**

Data mining is the process of analyzing large datasets to extract patterns, trends, and insights. It involves several stages, including data preparation, modeling, and evaluation. Data mining is used in many fields to improve decision-making, predict outcomes, and identify opportunities. The technique can be used to analyze customer behavior, identify fraud, optimize business processes, predict future trends, and more.

### **Task5**

#### **what are the Scrapping tools do companies use to collect data?**

Companies use various web scraping tools to collect data from websites. Here are some of the commonly used web scraping tools:

1. Beautiful Soup: Beautiful Soup is a Python library that can be used to parse HTML and XML documents. It allows users to extract data from HTML and XML files using a variety of methods.

2. Scrapy: Scrapy is a Python-based web scraping framework that allows users to extract data from websites. It provides tools for navigating and scraping websites, as well as for processing and storing the scraped data.
3. Selenium: Selenium is a web automation tool that can be used for web scraping. It allows users to automate web browsers, which can be useful for scraping websites that require user interaction.
4. Octoparse: Octoparse is a visual web scraping tool that allows users to extract data from websites without the need for programming. It provides a web-based interface for creating and executing web scraping tasks.
5. ParseHub: ParseHub is a web scraping tool that allows users to extract data from websites using a web-based interface. It provides tools for navigating and scraping websites, as well as for processing and storing the scraped data.
6. WebHarvy: WebHarvy is a visual web scraping tool that allows users to extract data from websites without the need for programming. It provides a point-and-click interface for selecting and scraping data from websites.
7. Apify: Apify is a web scraping and automation platform that provides tools for scraping websites, as well as for automating web tasks. It allows users to run scraping tasks in the cloud, which can be useful for processing large amounts of data.
8. Import.io: Import.io is a web-based platform that allows users to extract data from websites using a visual interface. It provides tools for navigating and scraping websites, as well as for processing and storing the scraped data.
9. WebSundew: WebSundew is a web scraping tool that allows users to extract data from websites without the need for programming. It provides a visual interface for selecting and scraping data from websites.
10. Content Grabber: Content Grabber is a web scraping tool that allows users to extract data from websites using a visual interface. It provides tools for navigating and scraping websites, as well as for processing and storing the scraped data.
11. Dexi.io: Dexi.io is a web scraping and data extraction platform that provides tools for scraping websites, as well as for processing and storing the scraped data. It allows users to run scraping tasks in the cloud, which can be useful for processing large amounts of data.
12. Web Scraper: Web Scraper is a Chrome extension that allows users to extract data from websites using a visual interface. It provides tools for navigating and scraping websites, as well as for processing and storing the scraped data.

13. OutWit Hub: OutWit Hub is a web scraping tool that allows users to extract data from websites using a visual interface. It provides tools for navigating and scraping websites, as well as for processing and storing the scraped data.

These are just a few more examples of the web scraping tools that companies use to collect data from websites. The choice of tool will depend on the specific requirements of the project and the expertise of the team performing the web scraping.