**Practical 1:**

**Aim: Write summarized description of any 10 data mining tools including Weka and Rapid Miner**

**Top 10 Data Mining Tools:**

1. **MonkeyLearn** | No-code text mining tools
2. **RapidMiner** | Drag and drop workflows or data mining in Python
3. **Oracle Data Mining** | Predictive data mining models
4. **IBM SPSS Modeler** | A predictive analytics platform for data scientists
5. **Weka** | Open-source software for data mining
6. **Knime** | Pre-built components for data mining projects
7. **H2O** | Open-source library offering data mining in Python
8. **Orange** | Open-source data mining toolbox
9. **Apache Mahout** | Ideal for complex and large-scale data mining
10. **SAS Enterprise Miner** | Solve business problems with data mining

**1. MonkeyLearn**

MonkeyLearn is a machine learning platform that specializes in text mining. Available in a user-friendly interface, you can easily integrate MonkeyLearn with your existing tools to perform data mining in real-time.

Start immediately with pre-trained text mining models like this sentiment analyzer, below, or build a customized solution to cater to more specific business needs.

MonkeyLearn supports various data mining tasks, from detecting topics, sentiment, and intent, to extracting keywords and named entities.

MonkeyLearn’s text mining tools are already being used to automate ticket tagging and routing in customer support, automatically detect negative feedback in social media, and deliver fine-grained insights that lead to better decision making.

With MonkeyLearn, you can also connect your analyzed data to MonkeyLearn Studio, a customizable data visualization dashboard that makes it even easier to detect trends and patterns in your data.

**2. RapidMiner**

RapidMiner is a free open-source data science platform that features hundreds of algorithms for data preparation, machine learning, deep learning, text mining, and predictive analytics.

Its drag-and-drop interface and pre-built models allow non-programmers to intuitively create predictive workflows for specific use cases, like fraud detection and customer churn. Meanwhile, programmers can take advantage of RapidMiner’s R and Python extensions to tailor their data mining.

Once you’ve created your workflows and analyzed your data, visualize your results in RapidMiner Studio, to help you spot patterns, outliers, and trends in your data.

Last but not least, this platform has a large and enthusiastic community of users, who are always on hand to help.

Try the free plan, which allows you to analyze up to 10,000 rows of data.

**3. Oracle Data Mining**

Oracle Data Mining is a component of Oracle Advanced Analytics that enables data analysts to build and implement predictive models. It contains several data mining algorithms for tasks like classification, regression, anomaly detection, predictions and more.

With Oracle Data Mining, you can build models that help you predict customer behavior, segment customer profiles, detect fraud, and identify the best prospects to target. Developers can use a Java API to integrate these models into business intelligence applications to help them discover new trends and patterns.

Oracle offers a 30-day free trial.

**4. IBM SPSS Modeler**

IBM SPSS Modeler is a data mining solution, which allows data scientists to speed up and visualize the data mining process. Even users with little or no programming experience can use advanced algorithms to build predictive models in a drag-and-drop interface.

With IBM’s SPSS Modeler, data science teams can import vast amounts of data from multiple sources and rearrange it to uncover trends and patterns. The standard version of this tool works with numerical data from spreadsheets and relational databases. To add text analytics capabilities, you need to install the premium version.

A 30-day free trial is available.

**5. Weka**

Weka is an open-source machine learning software with a vast collection of algorithms for data mining. It was developed by the University of Waikato, in New Zealand, and it’s written in JavaScript.

It supports different data mining tasks, like preprocessing, classification, regression, clustering, and visualization, in a graphical interface that makes it easy to use. For each of these tasks, Weka provides built-in machine learning algorithms which allow you to quickly test your ideas and deploy models without writing any code. To take full advantage of this, you need to have a sound knowledge of the different algorithms available so you can choose the right one for your particular use case.

Weka was originally designed to analyze data in the field of agriculture. Now, it is mainly used by researchers and industrial scientists, as well as for educational purposes. It is available to download for free under a GNU General Public License.

**6. KNIME**

KNIME is a free, open-source platform for data mining and machine learning. Its intuitive interface allows you to create end-to-end data science workflows, from modeling to production. And different pre-built components enable fast modeling without entering a single line of code.

A set of powerful extensions and integrations make KNIME a versatile and scalable platform to process complex types of data and use advanced algorithms.

With KNIME, data scientists can create applications and services for analytics or business intelligence. In the financial industry, for instance, common use cases include credit scoring, fraud detection, and credit risk assessment.

**7. H2O**

H2O is an open-source machine learning platform, which aims to make AI technology accessible to everyone. It supports the most common ML algorithms and offers Auto ML functions to help users build and deploy machine learning models in a fast and simple way, even if they are not experts.

H2O can be integrated through an API, available in all major programming languages, and uses distributed in-memory computing, which makes it ideal when analyzing huge datasets.

**8. Orange**

Orange is a free, open-source data science toolbox for developing, testing, and visualizing data mining workflows.

It is a component-based software, with a large collection of pre-built machine learning algorithms and text mining add-ons. It also has extended functionalities for bioinformaticians and molecular biologists.

Orange also allows for interactive data visualization, offering numerous graphics like silhouette plots and sieve diagrams, and non-programmers can perform data mining tasks through visual programming in a drag-and-drop interface. Developers, meanwhile, can opt to mine data in Python.

**9. Apache Mahout**

Apache Mahout is an open-source platform for creating scalable applications with machine learning. Its goal is to help data scientists or researchers implement their own algorithms.

Written in JavaScript and implemented on top of Apache Hadoop, this framework focuses on three main areas: recommender engines, clustering, and classification. It’s well-suited for complex, large-scale data mining projects involving huge amounts of data. In fact, it is used by some leading web companies, like LinkedIn or Yahoo.

Apache Mahout is free to use under the Apache licence and it’s supported by a large community of users.

**10. SAS Enterprise Mining**

SAS Enterprise Miner is an analytics and data management platform. Its goal is to simplify the data mining process to help analytics professionals turn large volumes of data into insights.

Through an interactive graphical user interface (GUI), users can generate data mining models fast, and use them to solve critical business issues. SAS provides a rich set of algorithms for preparing and exploring data, and for building advanced predictive and descriptive models.

Companies can use SAS Enterprise Mining for fraud detection, resource planning, and increase response rates on marketing campaigns, among other applications.

A free software trial and customized pricing packages are available.