**SparkML** is a package in **Apache Spark** that provides high-level APIs for creating and tuning machine learning pipelines. [These pipelines organize data preprocessing and modeling steps, making it easier to manage and deploy machine learning models1](https://towardsai.net/p/machine-learning/big-data-pipelines-with-sparkml)[2](https://dev.to/cloudmlguy/machine-learning-pipelines-with-spark-introductory-guide-part-1-1igl).

Here are **five free reference links** where you can learn more about SparkML:

1. [**Towards AI**: This article explains how to create Apache Spark ML pipelines for big data analysis, using accelerometer recordings from the HMP dataset as an example1](https://towardsai.net/p/machine-learning/big-data-pipelines-with-sparkml).
2. [**Dev.to**: A comprehensive guide to Spark machine learning pipelines, including an introductory part that covers the basics](https://towardsai.net/p/machine-learning/big-data-pipelines-with-sparkml)[2](https://dev.to/cloudmlguy/machine-learning-pipelines-with-spark-introductory-guide-part-1-1igl).
3. [**Apache Spark Documentation**: The official Spark ML programming guide, which provides details on SparkML’s features and usage](https://towardsai.net/p/machine-learning/big-data-pipelines-with-sparkml)[3](https://spark.apache.org/docs/1.2.2/ml-guide.html).
4. [**Coursera**: A course on machine learning with Apache Spark, covering topics like pipelines and data engineering](https://towardsai.net/p/machine-learning/big-data-pipelines-with-sparkml)[4](https://www.coursera.org/learn/machine-learning-with-apache-spark).
5. [**Kaggle**: Explore Kaggle’s resources on pipelines, cleaner code, and model validation1](https://towardsai.net/p/machine-learning/big-data-pipelines-with-sparkml).

Feel free to dive into these references to enhance your understanding of SparkML! 🚀