**H2O Integration**

**Features of H2O for integration**

* + 1. Data Loading
    2. Model Training
    3. Model Deployment

**Data Connectors**

All the above-mentioned features can be integrated into DS using **any** of the following methods:

1. Instance of H2O on Python/R/Java
2. REST API + H2O Instance
3. POJO (Plain Old Java Object) + JVM
4. H2O instance on YARN
5. Cloud Integration

**Supported File Formats**

H2O currently supports the following file types:

* CSV (delimited) files (including GZipped CSV)
* ORC
* SVMLight
* ARFF
* XLS (BIFF 8 only)
* XLSX (BIFF 8 only)
* Avro version 1.8.0 (without multifile parsing or column type modification)
* Parquet

**Default Data Sources**

* Local File System
* Remote File
* S3
* HDFS
* JDBC
* Hive

Relational databases that include a JDBC (Java database connectivity) driver can be used as the source of data for machine learning in H2O. Currently supported SQL databases are MySQL, PostgreSQL, MariaDB, Netezza, Amazon Redshift, and Hive.

Most other databases can be used using HDFS API calls.