**Framework used scrapy**:

* Scrapy is powerful framework for extracting data, processing it and then save it.
* It is fast and efficient this is because scrapy written with twisted asynchronous networking framework for python.
* Request are schedule and processed asynchronously. And Sending multiple concurrent request at a time.
* It has built in support for selecting and extracting data from HTML/XML using xpath and css selectors.
* It handles Robot.txt, user agent and scrapy has built in data pipeline like csv, Json etc.

**Components ->**

1. Scrapy engine : engine gets initial request to crawl from spider. It is responsible for controlling data flow between all components.
2. Scheduler : it receives request from engine and enqueues them.
3. Downloader : is responsible for fetching pages and feeding them to engine.
4. Spider : custom classes to parse responses and extract item from requests.
5. Item pipeline : processing item once they have been extracted and storing into the databases
6. Downloader middleware : it process request and give responses that pass from downloader.

**Database : Mysql**

**InnoDB:**

* InnoDB is a general-purpose storage engine that balances high reliability and high performance.
* Insert, update and deletes are optimized by an automatic mechanism.
* Create and drop indexes and perform other DDL operations with much less impact on performance and availability.

**Connectivity:**

* Install mysql-connector –python
* Create a connection using Mysql
* Create new database and table using SQL
* Write the code and scrap the data . show into mysql databse.