***Stop***

**For each Movie URL *U* in the input file**

**Get the movie text and release poster from Wikipedia**

***STEP-1:* *Get Movies URLs***

**(1) (2)**

**(3)**

**(4)**

**(1). Get the list of movie URLs for year *y***

**(2). Visit Wikipedia URL and scrape the list of Movies and their URLs**

**(3). Get the URL for year *y+1***

**(4). If all the URLs are done, store movies and their URLs as a CSV file**

**Get Movies URL for year *“y”***

**For each year *“y”* between 2000 and 2016**

***Start***

***Wikipedia***

***STEP-2: Scrape Movie plot and release poster* (2)**

**(1)**

**(4)**

**(3) (1). For each URL *U* in the file get movie text & poster**

**(5) (2). Visit Wikipedia URL and scrape movie plot & poster**

**(3). Save movie text and poster onto local disk**

**(4). Get the next URL**

**(5). If NO more URLs, STOP the process**

**Input**