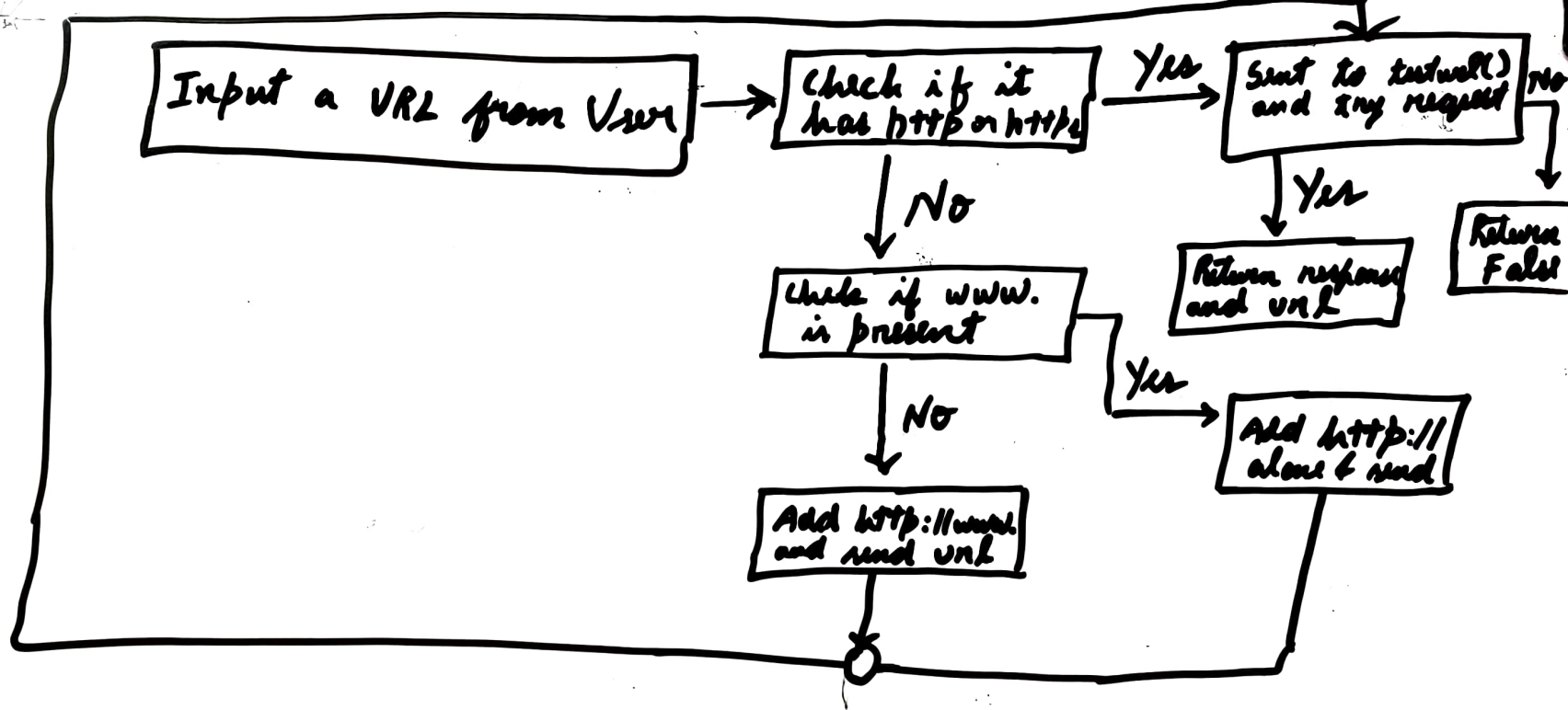


Block → 1



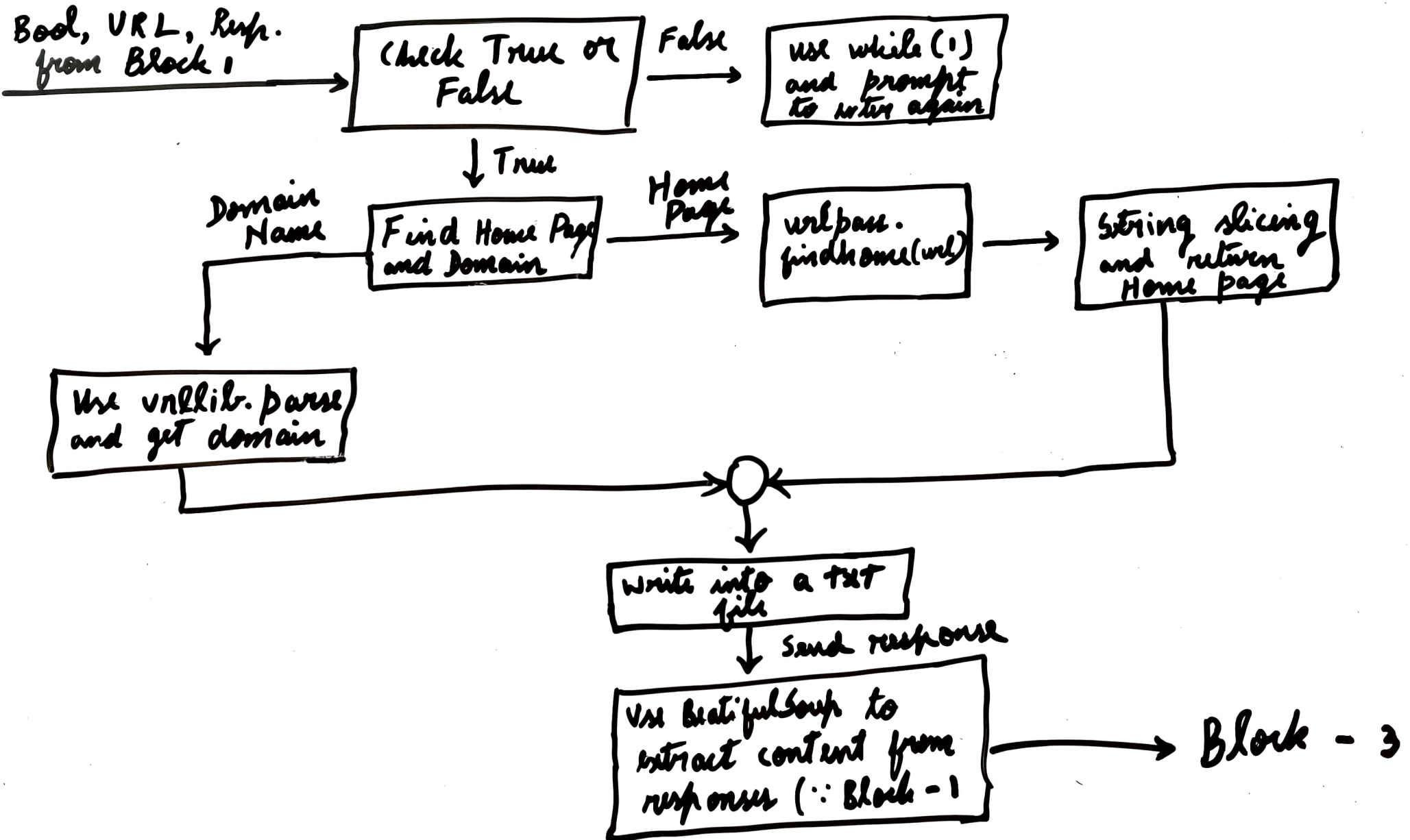
Checking if a URL is valid or not
using user-defined package "urlmate" → "urlpass" module which
has

1. isurl() method → check for "http://" and "www."
2. testurl() method → Make a request using requests module and return 3 Values.

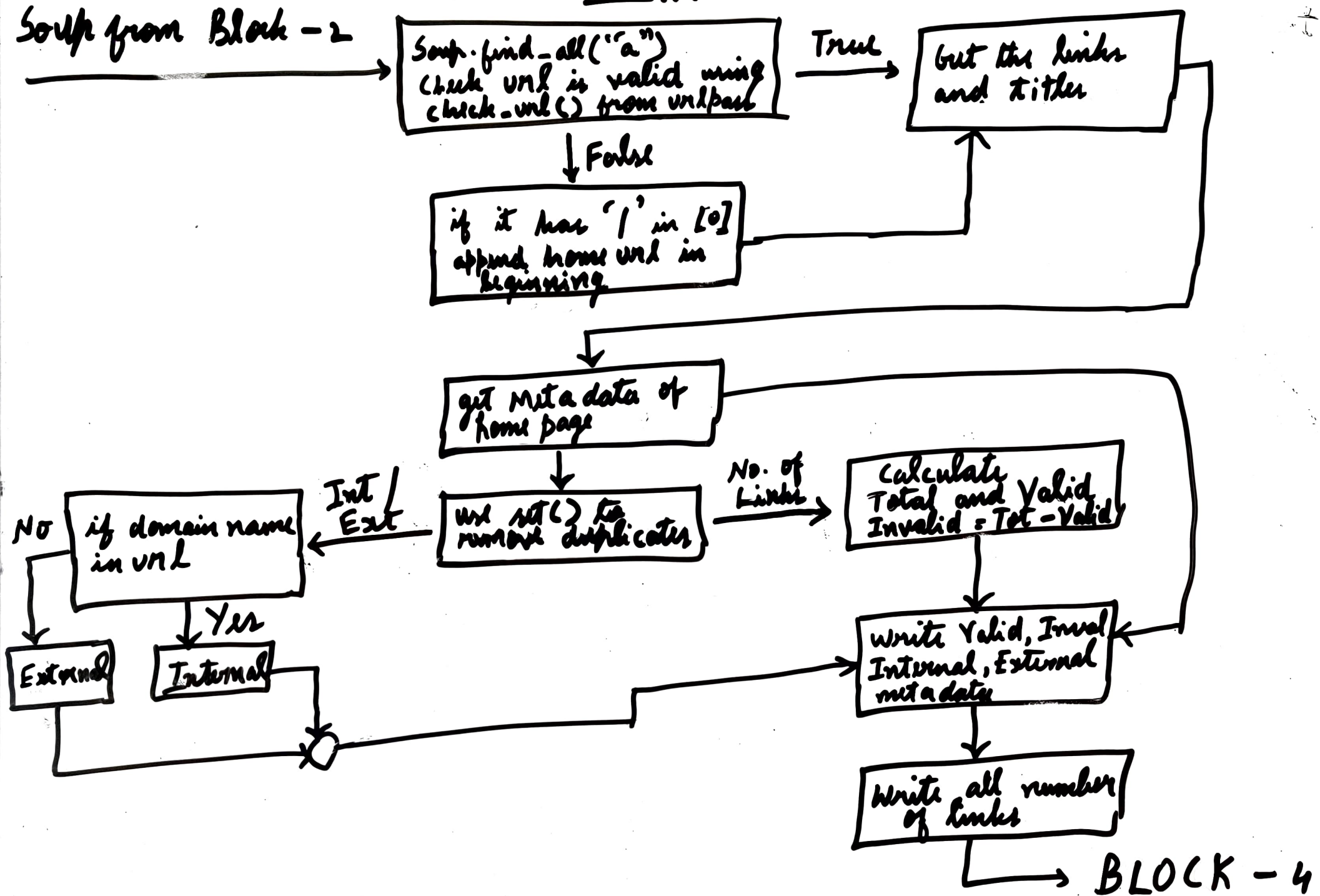
Values :

- * Bool
- * Complete URL
- * Response

BLOCK - 2



BLOCK → 3



BLOCK - 4

Valid-links, Home Page

Block - 3

text, unigrams, bigrams and
meta data from valid links
using get-text-from-url()
in wrtpass module

Do this for
Home Page

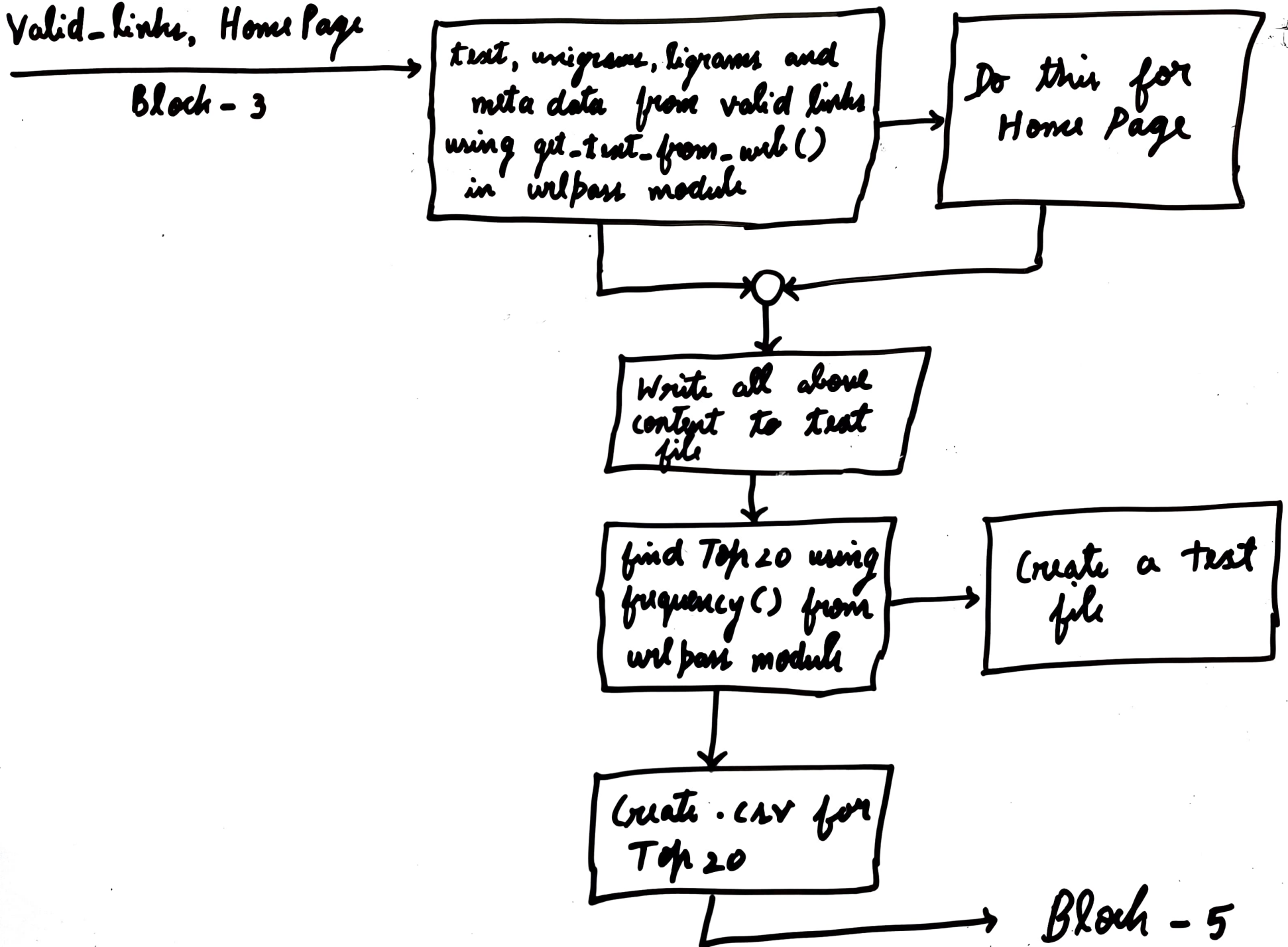
Write all above
content to test
file

find Top 20 using
frequency() from
wrtpass module

Create a test
file

Create .csv for
Top 20

Block - 5



BLOCK → 5

will pass `pagesize()` to
find the page size
of all links



Convert bytes to
KB



Store in Text
File