

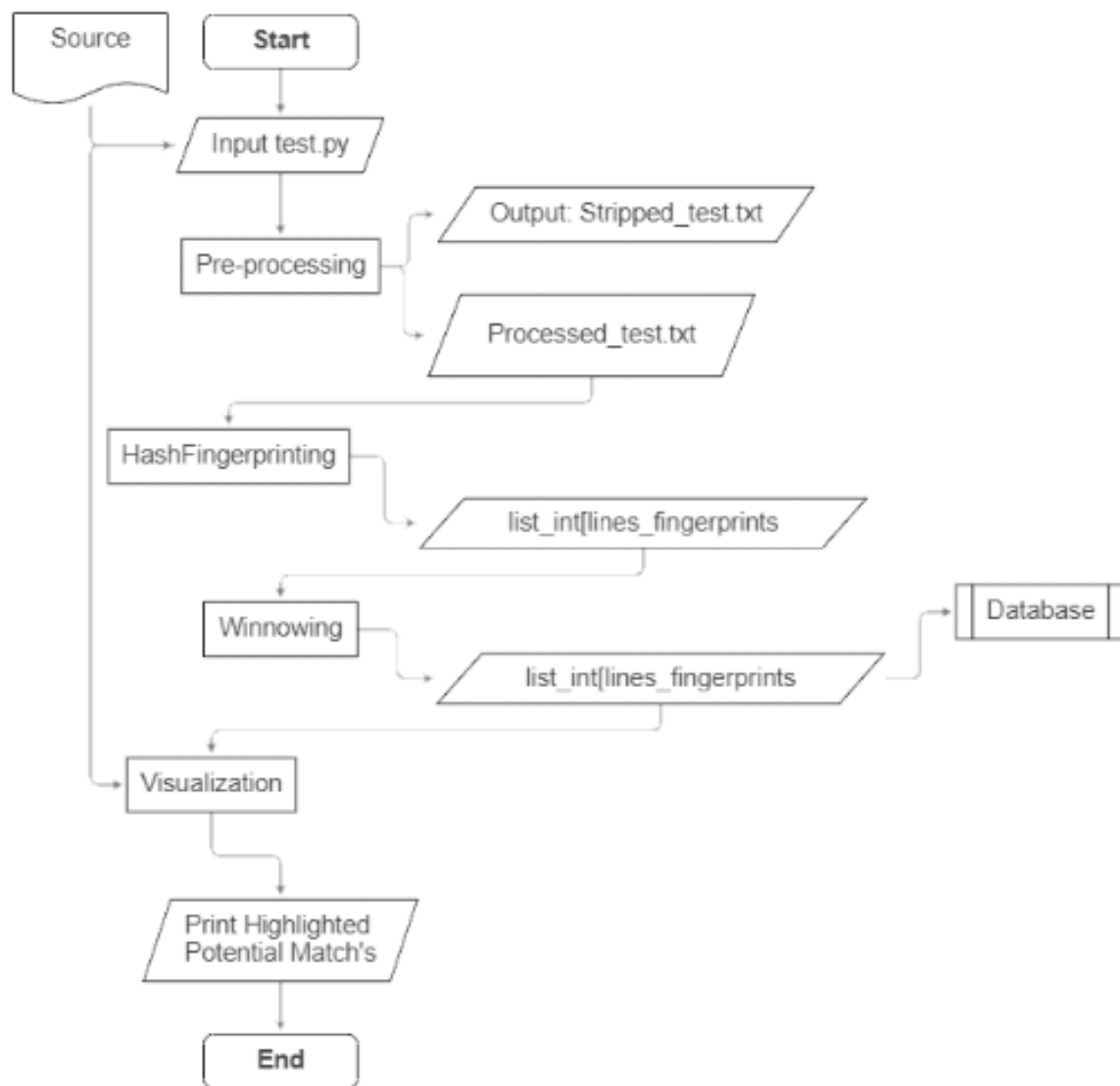
Preprocessing - Chase Jamieson

Hashing/Fingerprinting & Visual - Tracy Hotchkiss & Vinh Duong

Winnowing - Trevor Holland

# Plagiarism Detector Pipeline

## Pipeline Interface



## Visualization

source code:

```
# function definition and declaration
```

```
def calculate_sum(a,b):
```

```
    sum = a+b
```

```
    return sum
```

```
# The below statement is called function call
```

```
print(calculate_sum(2,3)) # 5
```

# Plagiarism Detector

Preprocessing	
Input: source file	Example ('test.py')
Output: stripped source file	Example ('stripped_test.txt')
Output: processed source file	Example ('processed_test.txt')
HashFingerprinting	
Input: processed source file	Example ('processed_test.txt')
Output: fingerprints source file	Example ('fingerprints_test.txt')
Winnowing	
Input: fingerprints source file	Example ('fingerprints_test.txt')
Output: list lines and fingerprints	Example (list_int[lines_fingerprints])
Visual	
Input: source file	Example ('test.py')
Input: list lines and fingerprints	Example (list_int[lines_fingerprints])
Output: Print Highlighted Potential Matches	