**Phase 1 Algorithm**

Step 1: Start

Step 2: **Take** Input Files

Step 3: **Extract** DNA Sequence of each files.

Step 4: Using String matching concept on each stored string in database.

Step 5: **Calculate** probability of matched strings from database.

Step 6: If probability **Greater Than** 7(assuming) then **Declare** Input file contain virus.

Step 7: Else if probability **Less Than** 7 **Declare** Input file partly contain virus or normal file.

Step 8: If step 7 is true **Go To** second phase.

Step 9: If step 6 is true **Stop** and **Declare** that the file is a Virus file.

**Phase 2 Algorithm**

Step 1: Start

Step 2: **Initialize** an **array1** contains symbolic form of, only first line of conventional malware signature.

Step 3: **Take** Input Files from first phase.

Step 4: **Cluster** input files according to their file formats.

Step 5: **Open** clustered files one by one and **Convert** each line of file into a symbolic form.

Step 6: **Initialize** an **array2** and **Store** all symbols of the file into the array.

Step 7: **Compare** array1 and array2.

Step 8: If array1 **Equal To** array2 then **Declare** Input file contain virus.

Step 9: Else Input file is virus free.

Step 10: Stop