Start: we chose the malware data-set for this project in order to determine what attributes are viable to distinguish between good and bad our attributes are as follow: ().

We used jupyter notebook and sci-kit learn to establish our criterias.