https://www.kaggle.com/datasets/ozlerhakan/spam-or-not-spam-dataset