

```
{ "nbformat": 4, "nbformat_minor": 0, "metadata": { "colab": { "provenance": [] },
"kernel_spec": { "name": "python3", "display_name": "Python 3" }, "language_info": {
"name": "python" } }, "cells": [ { "cell_type": "markdown", "source": [ "##IMPORT
LIBRARIES" ], "metadata": { "id": "LzQq1eTTHyi3" } }, { "cell_type": "code",
"execution_count": 1, "metadata": { "colab": { "base_uri": "https://localhost:8080/" }, "id":
"iA3hludMrP6u", "outputId": "8868bade-6029-4c9e-8c23-64424b2b24c1" }, "outputs": [ {
"output_type": "stream", "name": "stderr", "text": [ "[nltk_data] Downloading package
stopwords to /root/nltk_data...\n", "[nltk_data] Unzipping corpora/stopwords.zip.\n" ] },
"source": [ "import pandas as pd\n", "import numpy as np\n", "import nltk\n", "import re\n",
"\n", "nltk.download('stopwords')\n", "\n", "from nltk.corpus import stopwords\n", "from
nltk.stem.porter import PorterStemmer" ] }, { "cell_type": "markdown", "source": [
"##LOAD DATASET" ], "metadata": { "id": "_5ocKHanHwk5" } }, { "cell_type": "code",
"source": [ "a = pd.read_csv('/content/spam.csv',encoding='ISO-8859-1')\n", "a.head()" ],
"metadata": { "colab": { "base_uri": "https://localhost:8080/", "height": 206 }, "id":
"H89HzIp_0HLJ", "outputId": "b499f656-d75c-4b0c-e2f6-7b8332b4c109" },
"execution_count": 2, "outputs": [ { "output_type": "execute_result", "data": { "text/plain": [
" v1 v2 Unnamed: 2 \\\n", "0 ham Go until jurong point, crazy.. Available only ... NaN \n", "1
ham Ok lar... Joking wif u oni... NaN \n", "2 spam Free entry in 2 a wkly comp to win FA
Cup fina... NaN \n", "3 ham U dun say so early hor... U c already then say... NaN \n", "4 ham
Nah I don't think he goes to usf, he lives aro... NaN \n", "\n", " Unnamed: 3 Unnamed: 4 \n",
"0 NaN NaN \n", "1 NaN NaN \n", "2 NaN NaN \n", "3 NaN NaN \n", "4 NaN NaN " ],
"text/html": [ "\n", "
\n", "
\n", "

```