Euclidean Distance on Normalized Token Frequency