1. Compute the bigram based similarity between two strings “nikon” and “nike”.

Answer:

bigrams of nikon={ni, ik, ko, on}

bigrams of nike={ni, ik, ke}

sim(nikon, nike)=2/4=0.5

2. Compute the Jaccard similarity between the instance set of class “Book” and that of class “Textbook”.

book

Introduction to Algorithm

Textbook

Introduction to Algorithm

Red Sorghum

Data Structure

The Republic of Wine

Data Structure

Operation System

Answer: Jaccard\_sim(book, textbook)=2/5=0.4