

Algorithm Explanation

The aim of our project was to predict the crime on receipt of a distress call to the police so that the police can accordingly send a response team. The only information available to the police is the caller ID (location of caller). Each city is identified by a unique Identification number.

We have used the principle of a Bayesian Classifier.

The descriptions of the mapper and reducer functions are as follows:

Mapper

The main aim of the mapper function is to calculate $P(\text{city} / \text{crime})$.

Each input row that the mapper receives is for a particular crime. The row also contains the number of instances of the crime in each city ' cni ' as well as the total number of instances of the crime ' tc '. The mapper performs the division ' cni/tc ' for each city.

The output of the mapper is a <key,value> pair as follows:

Key: The city number

Value: Contains the name of the crime, $P(\text{city} / \text{crime})$, cni , tc all concatenated together as a single text.

Reducer

The main aim of the reducer is to calculate the $P(\text{crime}/\text{city})$.

The output of the mapper is received. Firstly, the total crimes of all types ' T ' are calculated by taking a sum of all ' tc '. Secondly, the total crimes occurring in a particular city ' CT ' is also calculated by taking a sum of all ' cni '. Next we use the following formula for deriving the result

$$P(\text{city}/\text{crime}) = [P(\text{crime}/\text{city}) * P(\text{city})] / p(\text{crime})$$

Where

$$P(\text{city}) = CT/T \text{ and } P(\text{crime}) = tc/T$$

The output is a <key,value> pair as follows:

Key: The city number

Value: Name of the most possible crime and its probability