IIT Madras

ONLINE DEGREE

Computational Thinking
Professor Madhavan Mukund
Department of Computer Science
Chennai Mathematical Institute
Professor G. Venkatesh
Indian Institute of Technology Madras
Behaviour Analysis and Predictions Using Data Collected for Similar Customer Problem

Professor Madhavan Mukund: So, this bottom-up identification that we are trying so we could look at different questions. So, let us look at for instance our earlier analysis of our customers in this shopping bills. So we are.

Professor G. Venkatesh: Try to find.

Professor Madhavan Mukund: We try to decide which customers are similar to each other.

Professor G. Venkatesh: Similar to each other. In terms of the.

Professor Madhavan Mukund: So, what we had done was we are kind of noted down in each category how many items appear in the bills of that customer

Professor G. Venkatesh: 9 items.

Professor Madhavan Mukund: 9 items.

Professor G. Venkatesh: One row.

Professor Madhavan Mukund: So, not the actual quantity but how many times that category appears like and then we also did some grouping so we combined everything that was.

Professor G. Venkatesh: Correct.

Professor Madhavan Mukund: Various things under utility and so on household utility. So, then we came up these kind of numbers like so we said that these are the top 6 customers we identified in terms of number of

Professor G. Venkatesh: Purchases

Professor Madhavan Mukund: Items bills that they have purchases and we said for example that our method not bought any apparel item but had 18 entries for food, 5 entries for. So, we are done this in one direction. We had looked at the shopping data for each customer and try to find

similarities differences some patterns with that. But by what we were saying before just like if you know somebody is marks maybe you can find out.

Professor G. Venkatesh: Who it is.

Professor Madhavan Mukund: Who it is.

Professor Madhavan Mukund: What pattern of marks are they done well in chemistry, I have done badly in physics they know it must be that student only. So, here again, maybe you could just say if this person's most frequent purchases.

Professor G. Venkatesh: Food

Professor Madhavan Mukund: Food and maybe second most is

Professor G. Venkatesh: Something.

Professor Madhavan Mukund: So, we could ask maybe the top 3 some.

Professor Madhavan Mukund: Per category.

Professor G. Venkatesh: Top three purchase per category.

Professor Madhavan Mukund: Top three categories and order which is the what is the item that type of items that this customer buys most often then what is the second most often and third most often.

Professor G. Venkatesh: And then from that.

Professor Madhavan Mukund: See if we can identify the customer.

Professor G. Venkatesh: So, decision tree you want to make.

Professor Madhavan Mukund: Similar decision tree, but now.

Professor G. Venkatesh: But this time you are not using this time you are not using the so we are.

Professor Madhavan Mukund: First we have categories.

Professor G. Venkatesh: Only categories.

Professor Madhavan Mukund: Categories

Professor G. Venkatesh: One only one item name of the category first, but you ordered the

Professor Madhavan Mukund: We ordering it based on quantities.

Professor G. Venkatesh: Quantities.

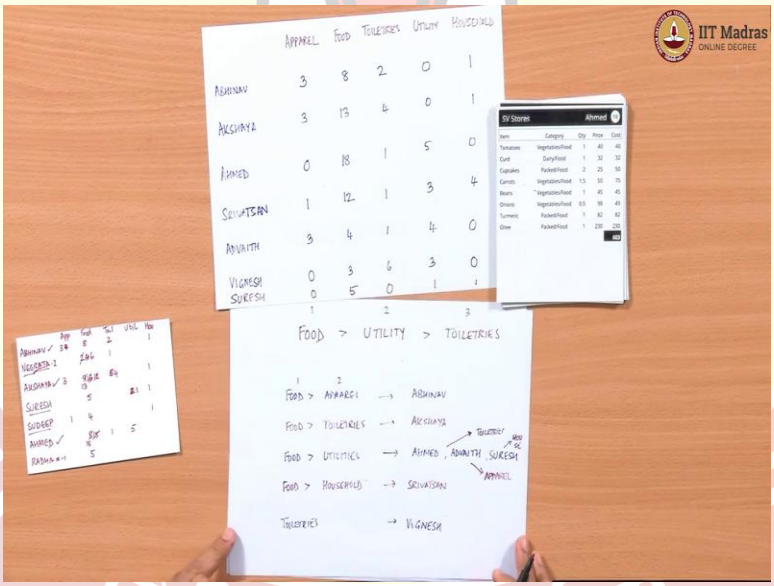
Professor Madhavan Mukund: But they are not asking the actual quantities, we just saying is it more than this is.

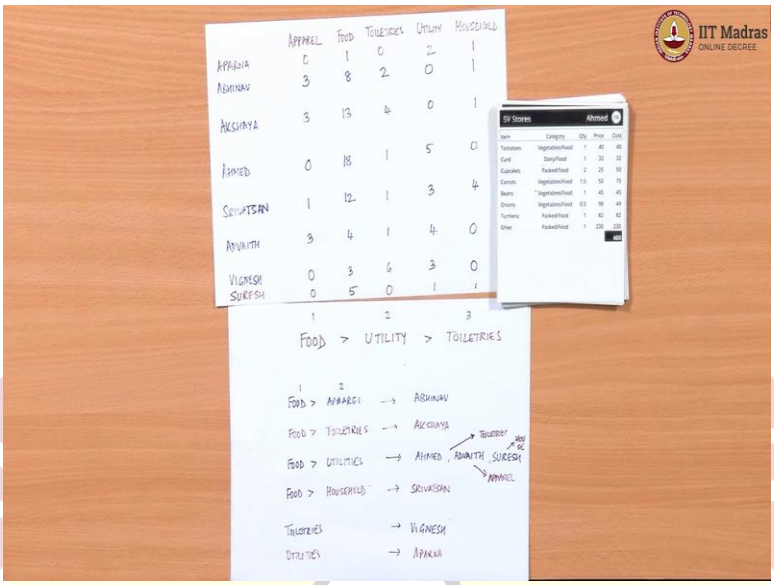
Professor G. Venkatesh: So, somebody bought highest amount their first they here are all the things they bought food is highest followed by apparel let us say clothing they about. Third thing they bought household.

Professor Madhavan Mukund: Correct.

Professor G. Venkatesh: Let us say,

(Refer Slide Time: 02:52)





Professor Madhavan Mukund: So, you want to ask.

Professor G. Venkatesh: Then that is that one person or more than one person like that…

Professor Madhavan Mukund: Something like if somebody's food is say bigger than say utility these are the top three.

Professor G. Venkatesh: Three things that the bought, food is greater than utility greater than.

Professor Madhavan Mukund: Toiletry say for example. So, if I have somebody who is number 1 so this is number 1, number 2, number 3 so it is not just greater than but they are the top three category items that this person is.

Professor G. Venkatesh: Who is this person?

Professor Madhavan Mukund: So, there is this uniquely identify the customer.

Professor G. Venkatesh: The only thing I am telling you basically is that he bought 3 items firstly he brought 3 items.

Professor Madhavan Mukund: 3 types of items

Professor G. Venkatesh: Most have brought all 3.

Professor Madhavan Mukund: Must have brought all 3. Otherwise, if it is 0 then I am not going to put it.

Professor G. Venkatesh: He has bought 3 items.

Professor Madhavan Mukund: And if I count the amount of each item that he has bought then these are the relatively relative.

Professor G. Venkatesh: That is how so first the out of the items he bought he bought 3 items and the out of the items he brought the maximum number of items was food second largest number of items if you took the shopping cards.

Professor Madhavan Mukund: Across all the cards

Professor G. Venkatesh: Cards

Professor Madhavan Mukund: All the all the bills of.

Professor G. Venkatesh: All the purchases he put it on just like a lifetime

Professor Madhavan Mukund: Behaviour

Professor G. Venkatesh: Life time analysis, put all of it in one card, then you see the food is the largest all of by utility for a followed by toiletries. Act is it enough to identify the person enough.

Professor Madhavan Mukund: Correct

Professor G. Venkatesh: Okay I now How do you do this is like so.

Professor Madhavan Mukund: we could be here so like a tree so we can.

Professor G. Venkatesh: You first check but in this case is food some other game at utility first.

Professor Madhavan Mukund: Correct, might be utility first, so we have to attach a label to each person. How do we do this?

Professor G. Venkatesh: Or you take branch each time 6 items, I do not know what you do, how you do that.

Professor Madhavan Mukund: So, I suppose what we could do is we could start with 2 and then we could order then this way or that way is food is bigger than utility food smaller than utility.

Professor G. Venkatesh: Like that such a tree.

Professor Madhavan Mukund: And then we could decide whether toilet is goes on top or.

Professor G. Venkatesh: Actually okay, so 3 but there are 6 items.

Professor Madhavan Mukund: So, we will have to insert C3.

Professor G. Venkatesh: 6P3

Professor Madhavan Mukund: 6p3 lot of lot of possibilities are there

Professor G. Venkatesh: 6 into 5 into 4 6 into 5 30 into 4 120.

Professor Madhavan Mukund: So, 5.

Professor G. Venkatesh: 120 enough.

Professor Madhavan Mukund: 5 are there.

Professor G. Venkatesh: 5 into 4 into 3.

Professor Madhavan Mukund: 60.

Professor G. Venkatesh: 5 into 4 20 into 3 60 60 different.

Professor Madhavan Mukund: Ways in which you can write this.

Professor G. Venkatesh: Write this 3. 60 you know we do not realize so that many are there. So, you want to list that this meaningless you do not want list all the 60.

Professor Madhavan Mukund: So, we want to find somehow which ones are there.

Professor G. Venkatesh: first take one pair at a time.

Professor Madhavan Mukund: So, we will try to incrementally build this. So, first of all we could.

Professor G. Venkatesh: We do it the way we did you do not algorithm.

Professor Madhavan Mukund: We will look at the data and then and so the data is in some sense we have got the data from these cards, but we have done ones round of processing the cards and we have summarized it in this kind of a tabular form. So, we have this new data and now in this new data we want to ask these questions.

Professor G. Venkatesh: Let us, start one by one. Let us, see what happens. We will go with Abhinav this the first guy on list. What is this his first item?

Professor Madhavan Mukund: Food

Professor G. Venkatesh: Second

Professor Madhavan Mukund: Apparel.

Professor G. Venkatesh: So, if food greater than apparel Abhinav, I am saying that is all do not need anything else.

Professor Madhavan Mukund: So, you have a rule which is food greater than apparel Abhinav.

Professor G. Venkatesh: Abhinav, let us, see we breaks will see now that time he will use his third thing which is toiletry or whatever.

Professor Madhavan Mukund: So, I come to the next row. Food is greater than apparel and his Akshaya.

Professor G. Venkatesh: But toiletry is there number 2 is toiletry.

Professor Madhavan Mukund; number 2 is toiletries. So, food is greater than toiletry.

Professor G. Venkatesh: This is 1 and this is 2.

Professor Madhavan Mukund: So, this is 1 and this is 2

Professor G. Venkatesh: So, in this case is food greater than toiletry. So, then it becomes Akshaya. So, far so good it is working.

Professor Madhavan Mukund: Food greater than utilities is Ahmed.

Professor G. Venkatesh: Food greater than utility than it is Ahmed. So, far so good. Then? Srivastsan food greater than household.

Professor Madhavan Mukund: Household

Professor G. Venkatesh: It work for this guy also. So, we only need 2. First 2 you take and that is done, looks like. But we are building it you looking at the data and figuring this out. That is the idea food greater than household.

Professor Madhavan Mukund: Shrivastsan.

Professor G. Venkatesh: Shrivatsan all right.

Professor Madhavan Mukund: Now, we have food and utility.

Professor G. Venkatesh: and utility

Professor Madhavan Mukund: Utility are the same. So, it would be like this and now we have a problem. So, we have Advaith also coming here.

Professor G. Venkatesh: Okay so.

Professor Madhavan Mukund: So, just 2 item is not enough.

Professor G. Venkatesh: Third item is what now for the.

Professor Madhavan Mukund: So, third item for Advaith is a apparel and for Ahmed it is toiletries.

Professor G. Venkatesh: Ok. so now you can separate them.

Professor Madhavan Mukund: So now you can separate them. So, if you go so Ahmed goes towards toiletries. If you ask the third question, I mean that you take the third item.

Professor G. Venkatesh: Only if you get food greater than utilities you ask this third question. Otherwise, you do not, okay then.

Professor Madhavan Mukund: And the last guy is.

Professor G. Venkatesh: Toiletries

Professor Madhavan Mukund: Toiletries is greater than.

Professor G. Venkatesh: So, in fact if the first item is toiletry I can say it is Vignesh. Do not even need to do anything. Do not you need this.

Professor Madhavan Mukund: Nobody else is in is a fan of buying toiletries first.

Professor G. Venkatesh: If it is toiletries, it must be Vignesh.

Professor Madhavan Mukund: So, these are like.

Professor G. Venkatesh: So, if the first item is toiletry declare it as Vignesh

Professor Madhavan Mukund: Vignesh.

Professor G. Venkatesh: The first item is food, then we cannot say.

Professor Madhavan Mukund: Then we have to ask at least one more item.

Professor G. Venkatesh: Second item is the second item is apparel.

Professor Madhavan Mukund: Abhinav.

Professor G. Venkatesh: Second item is toiletries.

Professor Madhavan Mukund: Akshaya.

Professor G. Venkatesh: But if it is utilities even second item is not enough third item is toiletry declare as it Ahmed it is apparel, declare as Ahmed. So, it works but we do not know it will work for him but other than this.

Professor Madhavan Mukund: So, now we have to look at all the other data and then combining.

Professor G. Venkatesh: Can take one more term just say, that even it works for or the.

Professor Madhavan Mukund: So, let me see we had written it down for some more people. Let us, see if we have. So, let me pick up Suresh. So, there is Suresh who is 5 food 0, 5 food, 0 toiletry, 1 utility and 1 household.

Professor G. Venkatesh: So, first item is food. So, you are here.

Professor Madhavan Mukund: Then second is say utility either you are in Ahmed question or Shrivatsan group depending on how you break this tie both are one. So, let us say it is utility. So, now we have Suresh also here,

Professor G. Venkatesh: And then.

Professor Madhavan Mukund: And now.

Professor G. Venkatesh: He has household so he separates.

Professor Madhavan Mukund: So, he is fortunately.

Professor G. Venkatesh: Lucky he got different one.

Professor Madhavan Mukund: Not bad, l

Professor G. Venkatesh: Not bad

Professor Madhavan Mukund: So, let us see if there is somebody else in this who is interesting from our earlier data that we had done. There is Aparna so maybe I will writer around top. So, these are all people who had few items so we did not include them in this top 6 but upper but Aparna 1, 2, 1. So, Aparna is the only person.

Professor G. Venkatesh: Who got utility.

Professor Madhavan Mukund: So, if I see utility

Professor G. Venkatesh: It must be Aparna

Professor Madhavan Mukund: As a top most thing.

Professor G. Venkatesh: So it works it, whether it seems to work see I mean if it does not work also, you can modify that is a cool thing about this method.

Professor Madhavan Mukund: So, we can change the we can refine the rule we can make a more elaborate rule like we had a rule with two things for Ahmed, when Ahmed came when Advaith came along we had to refine it to 3 and then when Suresh come along the 3 works and we got one more case for the third one. So, that is nice. So this kind of adapts itself, of course. It may have the order in which you get it you may kind of find different sequences maybe of. Now when somebody else comes to toiletries, then you will have to grow a second question for Vignesh and so on.

Professor G. Venkatesh: So, what you are saying basically if you give me a bill or give you three or four bills of one customer. I can look at those bills and I can say hey this must be Vignesh.

Professor Madhavan Mukund: And

Professor G. Venkatesh: By the behaviour

Professor Madhavan Mukund: By the behaviour.

Professor G. Venkatesh: Is what we do know about this very what Facebook and.

Professor Madhavan Mukund: But.

Professor G. Venkatesh: Guys are trying do that only right.

Professor Madhavan Mukund: But even the

Professor G. Venkatesh:

Professor Madhavan Mukund: But even the guys were selling you like Amazon or Flipkart.

Professor G. Venkatesh: they are also doing that.

Professor Madhavan Mukund: Because they are saying

Professor G. Venkatesh: what you buy.

Professor Madhavan Mukund: Because they will send you a promotional mails and you know you were interested in this kind of item to this this week it is on offer or no.

Professor G. Venkatesh: So, they are trying to figure you out by looking at what your bought.

Professor Madhavan Mukund: And what other people have bought also so they are also doing this similar customer business set we saw before they will also say people who bought this bought that saying you know why you do not also take a look at that encouraging you to look at products.

Professor G. Venkatesh: Combination of both first the identify who you are in some sense.

Professor Madhavan Mukund: So, they are trying to identify group in some sense not they do not specifically want you they want everybody who is like you and then try to sell things which are

Professor G. Venkatesh: Sympathy.

Professor Madhavan Mukund: By looking the or buying patterns they can target their so if you are the type of person who buys utilities you are unusual person like Aparna then you would send promotions.

Professor G. Venkatesh: Offer some

Professor Madhavan Mukund: Utility's to Aparna

Professor G. Venkatesh: And toiletries to Vignesh

Professor Madhavan Mukund: Toiletries to Vignesh

Professor G. Venkatesh:

Professor Madhavan Mukund: But food of course.