

Do you know your customers?

Presented by:

- Ghiles SIDI SAID
- Mohamed Ali DARGHOUTH
- Walid BELRHALLMIA





Overview

1. Problem description
2. Methodologies
3. Results

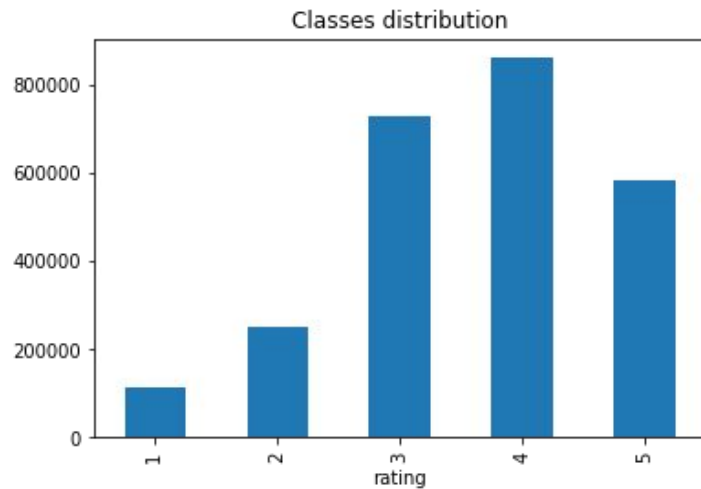


1. Problem description

A dataset about customer purchases and their reviews

3 columns: user, item and review score

2536704 rows, 92088 users, 3561 items.





2.Methodologies

From sequence classification to classification

Customer	Item	Rating
11676	Item3	2
11676	Item1	5
11676	Item2	3



Customer	Item1	Item2	Item3	Item4	Actual Item	Target
11676	0	0	0	0	3	2
11676	0	0	0	2	1	5
11676	5	0	0	2	2	3



RMSE using some basic model

Model	RMSE
Multinomial	1.5
Logistic Regression	1.14
Decision Tree	1.41



2.Methodologies

Matrix factorization

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	0	3	0	3	0
User 2	4	0	0	2	0
User 3	0	0	3	0	0
User 4	3	0	4	0	3
User 5	4	3	0	4	0

A matrix of user/item ratings

We see for example that user 1 has given item 2 a rating of 3.



Matrix factorization

	Feature 1	Feature 2
User 1	?	?
User 2	?	?
User 3	?	?
User 4	?	?
User 5	?	?

X

	Item 1	Item 2	Item 3	Item 4	Item 5
Feature 1	?	?	?	?	?
Feature 2	?	?	?	?	?

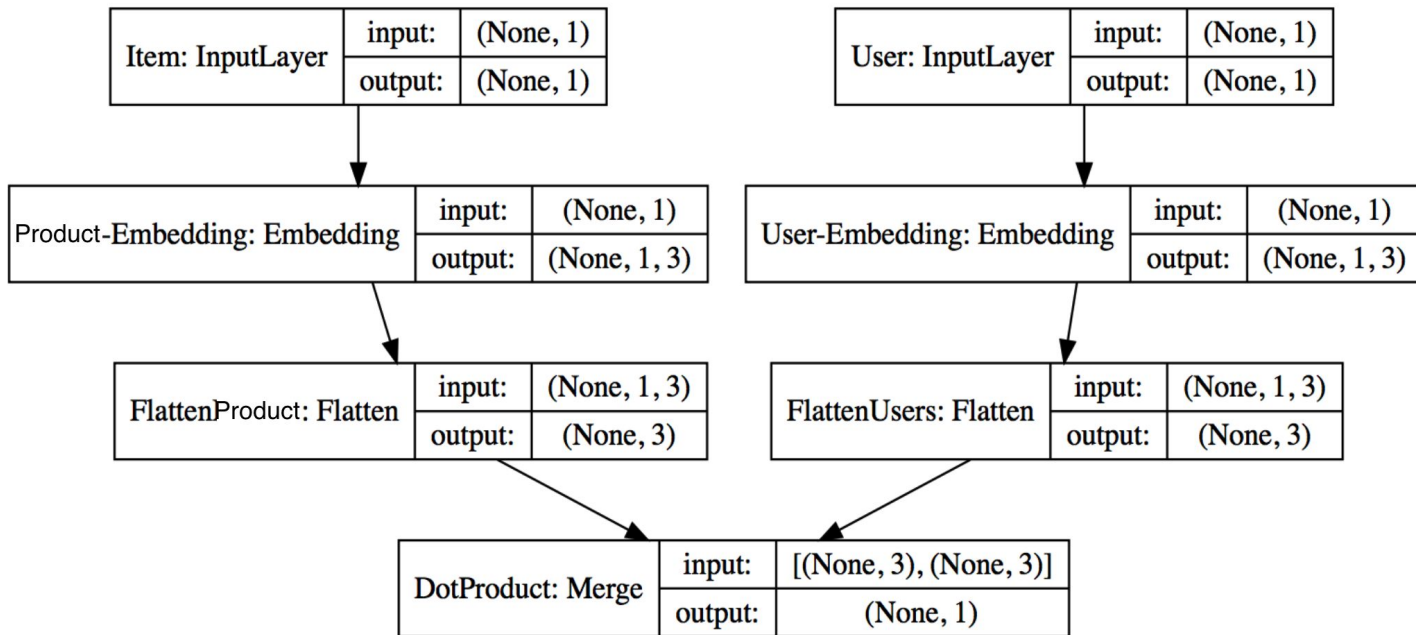
=

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	0.?	3	0.?	3	0.?
User 2	4	0.?	0.?	2	0.?
User 3	0.?	0.?	3	0.?	0.?
User 4	3	0.?	4	0.?	3
User 5	4	3	0.?	4	0.?

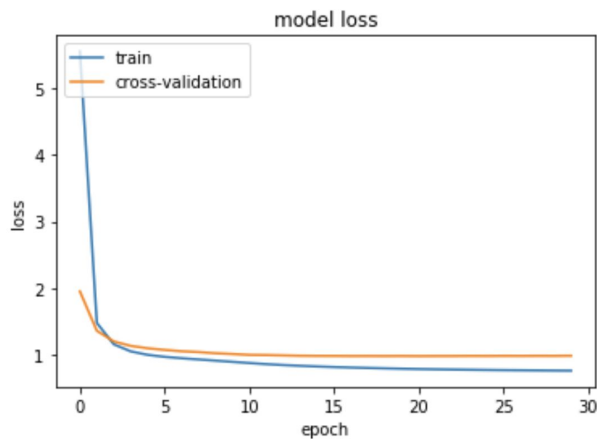


3.Results

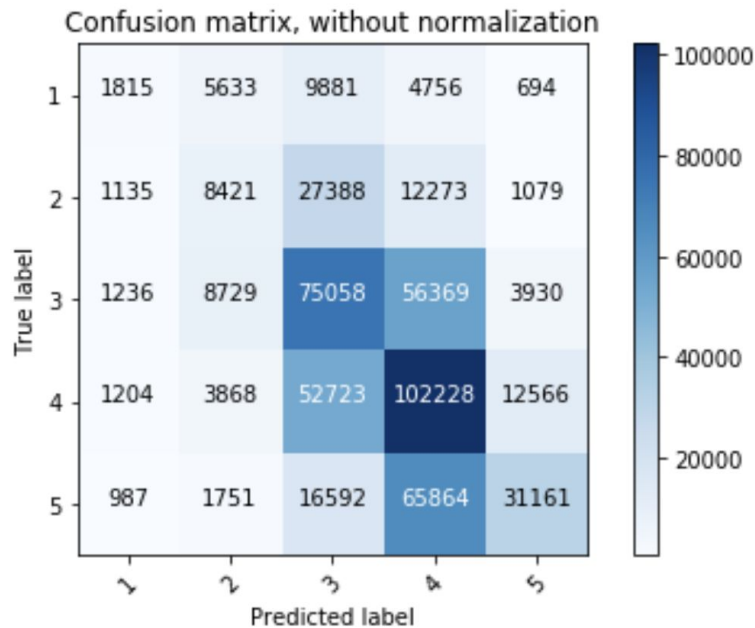
Architecture of the model using matrix factorization



Model using matrix factorization

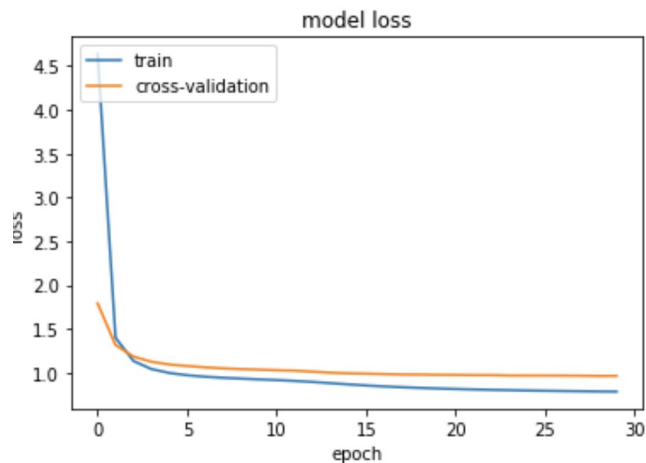


Model loss on Train and Validation Datasets

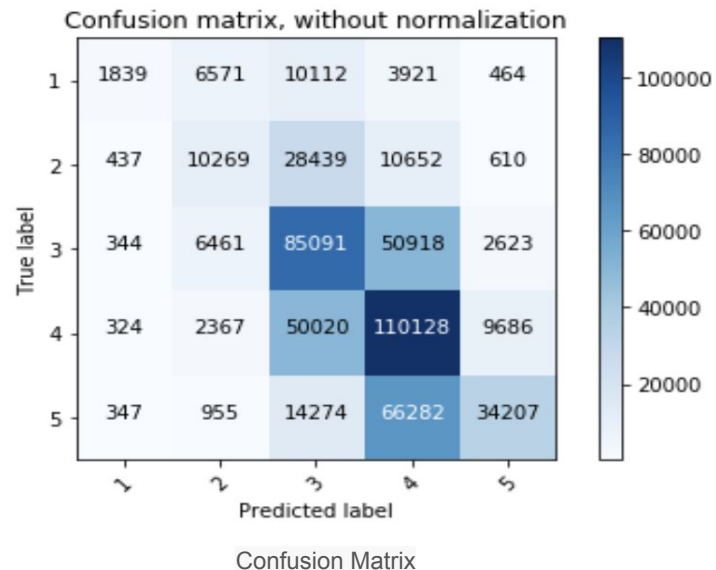


Confusion Matrix

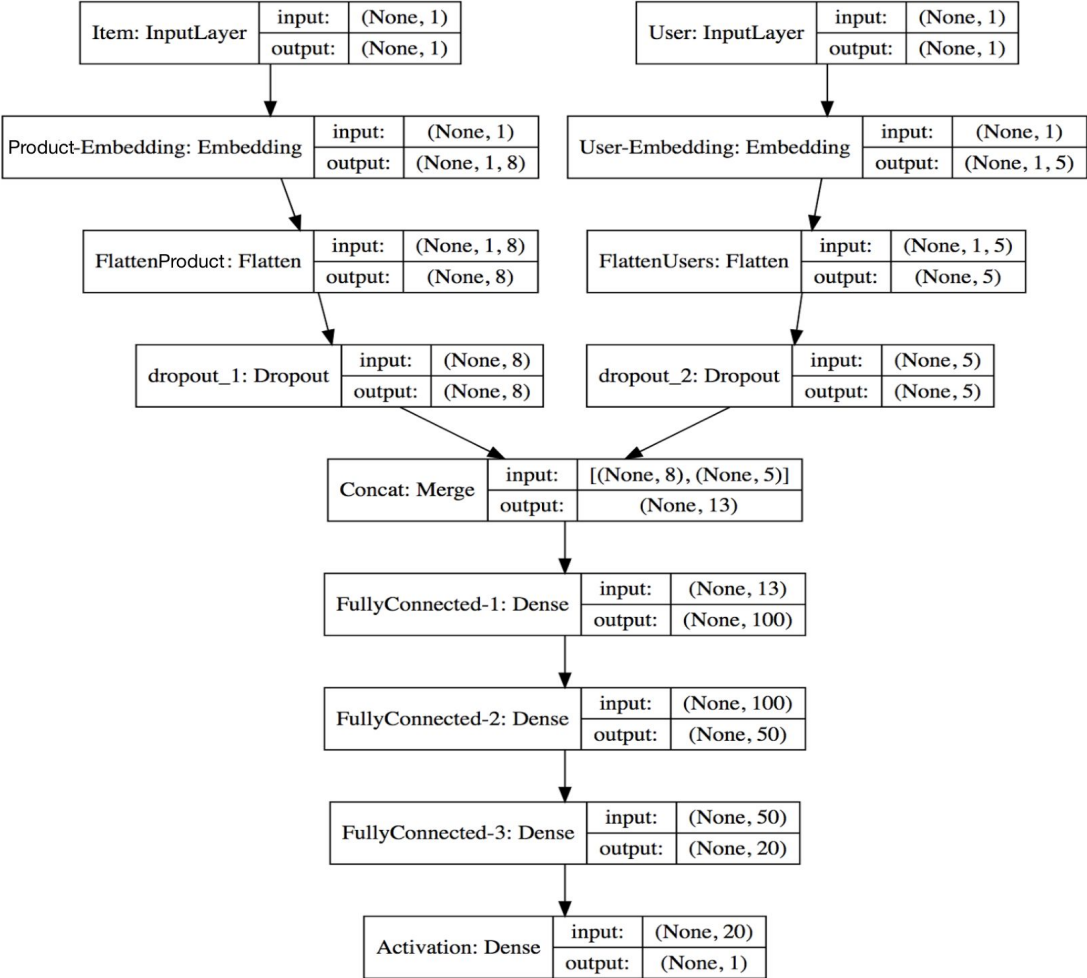
Model using non negative matrix factorization



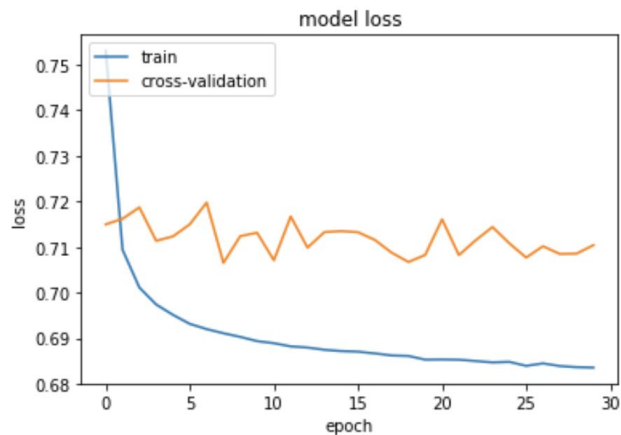
Model loss on Train and Validation



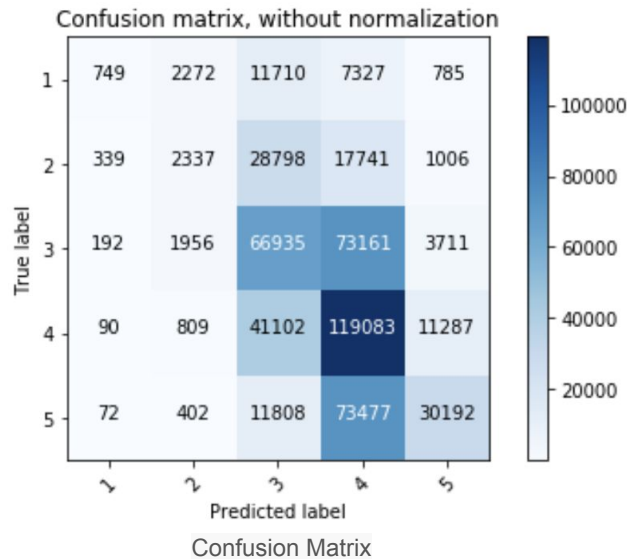
Model with merge



Model using merge



Model loss on Train and Validation Datasets





RMSE for the different model

Model	RMSE
Matrix factorization	0.932
Non negative matrix factorization	0.937
Merge based model	0.967