Topic Modelling of high and low ratings

In this section we use topic modelling to find common threads between reviews of score 5 and reviews of score 1, with the aim of discerning any patterns in the terms that contribute to the topic.

We explored several techniques for topic modelling namely NMF, SVD and LDA.

Non-negative matrix factorization

The NMF model decomposes its input matrix into two smaller approximate product matrices that only contain nonnegative values these are iteratively adjusted until they more closely result into the input matrix due to this process the features are clustered as the error value is reduce during each iteration.

NMF can take input matrices that have been processed by both term frequency and TF-IDF

Singular value decomposition

The SVD model decomposes the input matrix into its constituent parts in the form of 3 matrices. SVD acts as a feature reducer removing terms that are not important to the overall corpus.

Latent Dirichlet Allocation

LDA assumes that all topics follow a Dirichlet distribution across the documents in the corpus this leads to the probabilities of context between words being preserved. LDA groups together terms that occur together often into a topic which at times may not lead to topical grouping.

Observations

NMF TF

Score 5

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

not coffee tea food salt

one cup organic cat river

get starbucks teas cats murray

flavor vanilla white old store

would flavored numi canned small

eat cups spice no fingers

little taste flavor one time

want french orange wellness bowl

try blend green loves cooking

honey best black dry seems

found roast bags chicken table

since ground used eat food

work tully chai get one

think stars taste healthy use

still keurig wonderful needs amount

every not herbal weight sea

also go caffeine baby stuff

many drink get two salts

first coffees hibiscus time also

small cafe without feed keeps

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

great cookie product day shampoo

price cookies quality mph clear

product mother organic calories men

taste not cans speed scalp

tastes vanilla liver miles dandruff

amazon chocolate even peak not

snack amazon use exercise hair

easy good reviews drink ounce

make chip give work conditioner

years delicious best caffeine fluid

get taste skin avg feel

butter try stars celsius smell

loved creme amazon good therapy

tasting favorite company time one

really like would like like

chips company consistency not need

pack tea arrived works strong

year glad dog something head

right back control maybe mint

pasta used purchased mins shoulders

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

dog good like free love

dogs quality really gluten dogs

food well taste delicious buy

treats taste bread best cheese

variety price it chips powder

no flavor add pasta kids

price would real cookies absolutely

treat really best eat find

nature healthy cheese grams favorite

grain high tastes sugar little

small make soup tried drink

amazon made need soy milk

months bread much find my

free think beans ever cats

also products cook one healthy

buy pretty try would easy

picky cereal ham dairy awesome

dry wine make know keep

past fresh better amazon fruit

loves crackers pot bag tried

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

water ginger hot olive chocolate

coconut vernor sauce oil taste

tried soda flavor one bar

drink ale use oils bars

also use heat tried sugar

zico sugar frank full healthy

calories syrup also bodied dark

added michigan noodles give fat

best years not sure fruit

coco many perfect impressed better

vita amazon add way sweet

energy drink put different snack

buy since blend smooth whole

taste taste say low awesome

well corn tofu salads make

no gold beans somewhat tasting

brand cane tobasco turkey milk

little known nice turkish beans

sports green enjoy taste kind

way gnome really thought no

Score 1

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

not food product like tea

buy cat flavor tastes chai

could cats new flavor green

even ingredients beef try spice

money eat order tried not

get dry still tasted leaves

know meat use even flavor

company vet see lemon use

it bad spicy different blend

waste pet nong really much

anything also shim trying spices

store products packaging stuff teas

right fish said awful black

real dogs purchase get loose

sure natural page ever put

worth menadione made nothing box

either canned received drink tried

way time old something total

honey years get smelled return

disappointed including noodles bad weak

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

dog water coffee amazon taste

dogs coconut cup order better

problems natural even ordered bitter

poison concentrate cups item awful

treat plastic flavored get hot

chlorhexidine not mountain received smell

using added vanilla shipping drink

pee tried green package tasted

smell drink half disappointed first

vet brand dark customer tasting

not bottle price halva could

first zico drink return got

post made beans company honey

like tap caff purchase fruit

lick little grounds price brand

go get flavor opened mouth

used flavors instant box much

body one pods great something

long ever aroma sent artificial

chemical another bitter pack nasty

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

diet hair sugar made bag

science oil ingredient china popcorn

cats use artificial treats smell

tiki scalp drink chicken time

food acid calories dogs bags

evo bottle ingredients products microwave

bag greasy grams buy old

months sodium much jerky chips

always used stevia dog flavor

cat ci high usa since

get sensitive better not every

kitten using natural pet come

time product low sick burnt

weight tried healthy pets whole

pound natural per say strong

ago laureth aspartame another put

research clear free number open

not almond powder sweet pop

got sunflower contains us case

need might cereal read may

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

would good no one love

eat bad chocolate box tuna

cats really even bought eat

never candy cup time not

thought free donut buy us

money stuff way two vietnam

half gluten close cups stomach

great mix ever star wanted

even something money bad packed

since baking glazed yeast caught

caff tried give thought mercury

something pot name never really

use make think first thought

put tastes much get finish

pieces vitamins well better ounce

much looking flavor away california

maybe nothing might try wild

could foods little opened higher

green got bought even sardines

recommend time flavoring cup sandwich

When using NMF with a TF input there is a clear trend in the topics it is very easy to find an overarching commonality between the terms and possible “topic header” for example topic 9 of the score 5 group and topic 11 of the score 1 group are both about hair products. There is also a trend in the use of positive words in the Score 5 group and negative words in the Score 1 group, while the words No and Not appear in the Score 5 Group more telling words like bad, disappointment, weak and burn are seen in the Score 1 Group in comparison Score 5 group has no such words instead words like loves, like, quality and favourite are seen but some positive terms are seen in Score 1’s topics.

NMF IDF

Score 5

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

not coffee tea food product

good cup green cat great

like bold teas cats excellent

one great bitter old arrived

flavor strong white canned use

taste cups drink chicken taste

really decaf iced dry time

would dark black wellness years

get starbucks wonderful loves condition

use taste flavor feeding really

also blend great no service

eat good chai wet quickly

tried mountain bags feed fast

much flavor peach weight quality

delicious drink smooth problems smooth

well roast caffeine diet substitue

sweet maker loved month easy

little like loose vet cappuccino

make flavored peppermint needs buy

every favorite numi eats using

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

treats free chips sauce love

dogs gluten potato hot kids

training cookies bag heat drink

smell know flavor goes dogs

liver eat case say absolutely

maltese gf chip great my

give delicious mmmm anything favorite

freeze pasta ordering wings buy

small dairy stock frank cats

love granola fat try yummy

dried would whole tried bold

no rocks sons enjoyed energy

easy snacks family garlic time

size products addictive buy taste

three oatmeal crunch hooked milk

biscuit soy eating looked makes

perfect tasting alternative back little

pet muffins store given sweet

dog great madhouse awesome every

salmon disease pop sauces corgis

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

bars amazon chocolate best coconut

kind stores dark ever water

bar find milk the zico

fruit shipping hot mix electrolytes

candy local bar tasted refreshing

tasting glad delicious tried tastes

yum order sugar around sweet

healthy delicious company order sports

more found rich fast drink

order com again tasting potassium

blueberry grocery chocolates hcg added

nut high taste family brandsf

natural buy coffeemate makes thailand

flavors ordered creamer it nirvana

almond free cocoa brand brand

good each bit vanilla vita

larabars available error decaf mango

thinkthin difficult carrying pasta pure

cherry supermarket turbinado delivery calories

going thank sweetness healthiest coco

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

snack price dog peanut oil

healthy great loves butter olive

kids good treat skippy tuna

great high food great quality

quick buy my zuke oils

breakfast buying teeth trusted tins

tasty outstanding treats powdered cost

protein choking stomach delivered wonderful

crunchy though bathroom low taste

popchips even them cookies fresh

eating bulk pill around canola

fat trader little rawhide flash

easy wow gets name avocado

make happy sensitive reeses tried

sweet joe goes go several

satisfying case small too rice

bag picky jerky salmon acids

eaters sky perfect peanutty fatty

happy pack last tried monounsaturated

daughter hello likes purposes high

Score 1

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

not coffee cup amazon tea

like flavored cups product chai

taste weak explode order green

product beans keurig ordered leaves

would vanilla mountain shipping stash

buy aroma instant received spice

one bitter filter package teas

flavor grounds green beware spices

good strong box item black

even price again return flavor

tastes pods many refund iced

could starbucks kcups box weak

no instant reason date like

tried flavor hot expired flavors

bad ground brand disappointed tastes

money wolfgang grounds description seasonings

get roast one arrived celestial

bought dark coffees never strong

tasted raspberry twinings get lemon

try club alot customer honey

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

china coconut chocolate sugar cookies

made water old grams pamela

treats plastic raspberry healthy cookie

usa concentrate hot calories ginger

chicken natural melted added butt

products zico bar mango pamelas

dogs added white high bakery

pet bottle fresh tsp snap

dingo taste cacao artificial day

not pure dark low cases

read brand where much exact

pets drinking bloomed stevia version

trust milk favorite ingredient pain

no straight disappointed quaker found

sick come the coke oatmeal

dog real taste oatmeal according

buying close gritty taste restrictions

many made curry no reduction

fda coco neither you purchased

buy juice sticks carbs apparently

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

dog food stale cat bars

treat cats old touch box

dogs eat time treats chalky

treats would vendor not bar

toy dry sitting sniffs mango

ball pet line eat arrived

not ingredients rancid hates macadamia

refuses one purchase loves eating

touch vet guess maybe almond

eat day purchased near clif

chew dogs long my one

would cat disappointmentsticks old

meat feeding spices kitty think

bone sick fresh shelter larabar

vet bad badly oh kind

my old past extreme eat

it picky gave dry vosges

thing eating big times pie

wellness science ground buy pecan

lick diet now particular dipped

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

chips bag drink awful gluten

kettle popcorn drinks based free

brand bags sip strawberries semolina

chip pop ginger th barley

protein butter bad taste not

completely time lemon teas celiac

rancid covered artificial tasting discription

oil popped tasting brands baking

sour microwave iced love member

try lb actually reviews wheat

salt overpriced milk item ingredient

cream store glass purchased disease

chance plastic bitter sprayed mix

onion every taste smelling know

apple pet mixes cologne list

online pound drinking mouth search

style trash beer medicinal malt

salty full shampoo away chex

buying sealed immediately believe family

parmesan michaels coco thought that

The results of NMF IDF are just as good as those seen in NMF TF.

SVD

Score 5

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

not coffee tea salt salt

like tea organic food great

great cup teas coffee product

good taste white tea use

coffee flavor numi cat store

taste starbucks spice cup murray

one vanilla flavor dog river

tea flavored orange river taste

product teas bags murray fingers

food cups black dogs cooking

love great used no small

flavor drink green organic table

no green ginger cats time

get blend chai fingers good

would numi herbal small amount

also french hibiscus time bowl

best white get bowl stuff

tried like mu seems seems

amazon best lime store sea

use black wonderful also chips

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

product good good love great

great cookie product like day

dog taste organic day good

coffee cookies quality mph food

price chocolate day drink mph

love love mph dogs shampoo

food delicious cookie calories cat

quality food liver water hair

tea mother cans dog like

organic vanilla calories speed scalp

dogs free speed work men

amazon bars company coconut calories

free healthy like treats clear

liver bar reviews exercise speed

cans really stars peak drink

cup sugar miles miles dandruff

stars amazon exercise use exercise

even eat numi caffeine peak

best chip cookies avg miles

give gluten peak celsius conditioner

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

dogs taste like like love

dog organic hot free flavor

cookie like sauce gluten product

shampoo shampoo love ginger shampoo

cookies sugar good use chocolate

like bars use sauce free

love one flavor vernor one

clear bar cookie cookies day

men chocolate product cookie clear

scalp best cats hot best

great men food soda men

treats scalp cheese ale scalp

dandruff clear add amazon gluten

hair coconut cat best dandruff

feel no taste vanilla hair

amazon dandruff heat food milk

mother white beans used ounce

conditioner water cook sugar chips

fluid ounce quality delicious would

try healthy powder pasta sugar

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

water use love one chocolate

coconut ginger organic olive hot

chocolate vernor cats oil sauce

cookie sugar great flavor bar

drink soda spice sauce dark

cookies ale white hot beans

vanilla taste not no make

way years day tried one

one many cat good dog

amazon skin one use bars

mother treats numi organic day

energy no cookies dog heat

also syrup cookie full great

tried michigan eat chips perfect

zico still used best quality

little healthy orange oils granola

beans since ginger always noodles

no get mph ginger fat

get size many give kind

back dog olive treats oil

Score 1

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

not food product like tea

product dog amazon tea chai

like product order product not

taste cat new taste food

would cats made flavor spice

one diet received good green

food amazon ordered would leaves

no science still one amazon

good time beef sugar use

even dogs use no return

tea one purchase food blend

buy made item tastes ingredients

water eat spicy much box

dog order nong green spices

get ingredients shim water loose

could vet packaging bad total

flavor meat page try teas

much pet disappointed tried quantity

coffee chicken buy cats cat

bad products china time label

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

dog water water product would

like coconut dog would sugar

dogs taste coffee cats taste

made natural coconut like one

china sugar natural cat eat

treats product made coffee better

product added would diet artificial

problems diet tea use good

treat concentrate dogs science honey

poison food sugar eat calories

chlorhexidine science use not love

smell ingredients cup dry free

using plastic china food tuna

sick cat ingredients flavor could

chicken drink ingredient lemon dog

pee bottle drink caff low

used tried treats hair much

lick also green beef grams

chemical zico many animals ingredient

look noodles problems ingredients thought

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

diet one sugar made bag

taste would no china would

science good ingredients one smell

dog cats ingredient buy no

sugar diet natural treats popcorn

no science chocolate love taste

bag get like would flavor

good box made bag dog

tiki hair even chicken food

evo time list eat money

corn using products tuna water

always even china flavor opened

better use chicken say put

got never way sauce brand

months could drink products chips

kitten oil high hot never

long noodles hair box ingredients

calories put use dogs another

pound used contains noodles great

need size much never price

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

no good amazon many chocolate

one no tuna first bad

food love love used no

product tuna even ingredients even

flavor eat no natural cats

tuna flavor get also got

cup hair cup love box

love oil taste use honey

try really diet much taste

dog us science people cup

meat use green box much

yeast scalp us taste cat

ever stuff like using bought

tried vietnam brand ingredient green

make chicken thought order never

even packed wanted popcorn well

eat bad vietnam time time

really tried well tuna two

sugar wanted hair coffee would

bought little could bad money

SVD’s topics tends to have several improperly assigned terms in the Score 5 group topic 16 which seems to cover to topic of beverages has a out of place term “dog” this improperly assigned terms occurs in all topics with some topics having no clear topic.

LDA

Score 5

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

oil cheerios pouch tropical popcorn

olive tree pouches cytomax molasses

tins tassimo pills fruit pops

oils marshmallows donuts moles theater

virgin bunny dunkin hole packaging

zoe baja kittens tunnel honey

jug bonsai popper traps potassium

massive disc warm mole daughter

having pb baby set poper

ummmm uncooked pillpockets light ride

yourself kosher andy lindt tbsp

taster shape senseo black consume

decadence cooker granules trap chews

lark walmart spreading dirt sports

bertolli make expiration dug style

slimey love decided passion nutritional

velvety take my proof makes

roses dyed licked able popper

exceptional peeps open juice styrofoam

accepted tempting ordeal mounted tasteless

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

focus dog bars shortening ginger

jittery treat best spectrum vernor

engineering greenies bar basil ale

majoring vinegar like luscious soda

shots toy ever holy beer

mechanical momo sauce gobbles michigan

crutches capsule not unknowing upset

studying end noodles shoprite spicier

exams pet chocolate classmates gnome

burst gelatin one crisco isbn

metabolism bone chips allergan cane

energy minutes calories talking requested

beneficial seniors kind previous promptly

offer company mustard wegmans zest

staple canine tasting members drops

feeling sage garlic sorry exactly

breakfast terriers protein flavored delivered

benefits pectin mph butter accounting

home knee honey family page

morning pain snack fillers microbrews

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

treats creamer not bread hair

remember pump great baby tea

tangerine bottle product bubbles used

actually lid good toast day

beverage lift love biscuits echinacea

crisp pushing like gras anti

hips turning food boxes see

race bailey one foie organic

hansen voila taste loaf slim

hansens illy would breakage medicine

wax energized amazon babycook pantry

carbonated insomniac no create take

cable jumpy flavor box white

deserve nirvana also lightly weight

film cranberry get cooked months

creepy ito free device years

movie polluting delicious top almost

cheer brandsf buy broken remedy

possessed en price previous helped

character thailand dog de symptoms

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

coffee pasta mg cat shampoo

tea hooves per hot clear

not quinoa gm food scalp

like shopped red sauce men

cup structure kettle cats dandruff

great al black old conditioner

taste dente onion wellness fluid

flavor celiac sour heat therapy

love based spheres chicken shoulders

drink disease noticeable canned clean

water prague gelatinous subscription controlling

chocolate peat raspberries flavor separate

best wargs mono dry tree

good boggy bakes time itchy

one trapsing iron texas eucalyptus

tried scotch sliced not cooling

cups lounging represent feed pyrithione

coconut peaty center rub sudsy

favorite cologne german goes therapeutic

get hordes barely eats lasting

Score 1

topic 0 topic 1 topic 2 topic 3 topic 4

-------- -------- -------- -------- --------

product not not not not

not product product tea money

would taste even like buy

dog like date taste taste

no water cup water product

half no good flavor waste

post coconut expired would like

taste good like sugar no

review order no good would

one one expiration product box

caff get coffee green love

pee even natural no tea

dogs bought disappointed free one

buy much yeast mix hot

even could one tried bought

coffee made many tastes stuff

green price food even bag

mountain candy amazon drink get

read flavor get chai good

better amazon ordered buy horrible

topic 5 topic 6 topic 7 topic 8 topic 9

-------- -------- -------- -------- --------

one not product coffee not

bad product not not taste

not like stale product like

cups bag amazon amazon much

pizza no like cup would

raw food chocolate buy get

artificial taste old taste flavor

love one one even drink

review even order two water

bunnies get box like food

diet much love no tried

product would would one product

thought diet tastes disappointed good

amazon old arrived flavor popcorn

food eat good cups love

purchased amazon time bad first

terrible buy flavor never price

reviews order taste would licorice

flavors know ever pods try

really science tried received hard

topic 10 topic 11 topic 12 topic 13 topic 14

-------- -------- -------- -------- --------

not not food chocolate not

like made not not dog

product china dog sugar product

amazon would cat msg amazon

taste like cats bars would

one dog eat much shipping

no treats would get package

tried one one salt could

try buy dogs litter even

cookies no bad regular store

get know time raspberry made

food dogs like line one

good could chicken buying like

even product good would treats

really products even product no

bad say product really bottle

chips give meat better get

know better first healthy food

bought tried ingredients grams old

day something no like sweet

topic 15 topic 16 topic 17 topic 18 topic 19

-------- -------- -------- -------- --------

not not not not like

like like no would not

one taste would one tastes

box sugar like say cup

product would item buy taste

taste artificial brand taste cups

time good beans like good

really honey way hot sweet

flavor product product bad flavor

good tastes buy never one

peanut much free many even

star bad try made coffee

opened drink get could really

bad could box two boxes

tried one much tabasco product

amazon smell received first could

ordered awful coffee even tasting

butter stuff small thought box

going also open get horrible

tasted never shipping type pretty

LDA has mixed results there is no clear positive or negative arrangement of terms in the Score groupings. Some topics like Score 5 group topic 0 is clearly about Oil while theory topics like topic 0 of Score 1 Group has many unrelated terms in them making topic assignment unclear.

The NMF model has the most coherent results all topics have a clear theme and the Score groups do not have any contamination of positive and negative terms. LDA picks up word groups that occur commonly with one another leading to mis grouped terms while SVC approximates the groups very loosely.