

# Word Embeddings Result

Based on lyrics data

# Details

- Object : 1.2M lyrics
- Size : 1.3GB
- Form : 한글 Korean char, Alphabet (almost 50:50)
- Extract nouns via mecab, nltk libraries
- Eliminate stopwords
- Comparison among FastText, GloVe, Swivel, LSA

# 희망

## FastText

소망  
넘실거린다  
끝도  
꿈  
소망일  
묵호  
얏타맨  
주어짐  
젊은  
일의  
해방구  
외항선  
절망  
펼쳐갈  
창연  
버팀  
이나  
은해  
용솟음친다  
용달차

## FastText Jamo

maintains  
navigators  
enemies  
netflix  
navigate  
팔짱  
mofuckas  
dopin  
rookies  
invadas  
clouds  
umbrellas  
focus  
navigated  
seats  
locus  
audiences  
havik  
lids  
maintainin

## GloVe

꿈  
소망  
미래  
열정  
새로운  
란  
삶  
기쁨  
믿음  
라는  
승리  
작  
행복  
끝  
이란  
두려움  
의  
가득  
세상  
바람

## Swivel

찬란  
꿈  
세상  
소중  
행복  
삶  
인생  
열정  
절망  
소망  
바람  
이란  
도전  
깃발  
끝  
젊음  
힘찬  
나날  
라는  
청춘

## LSA

성  
고통  
상대방  
부드러움  
화살  
명암  
쓴맛  
공백  
중력  
명성  
앞날  
권력  
명분  
자부심  
치욕  
사상  
여운  
염증  
두근거림  
고난

## LSA + PPMI

성  
고통  
상대방  
부드러움  
화살  
명암  
쓴맛  
중력  
공백  
명성  
앞날  
권력  
명분  
자부심  
치욕  
사상  
여운  
두근거림  
염증  
고난

# 절망

## FastText

좌절  
고통  
슬픔  
두려움  
자괴  
무기력감  
때에  
슬픔  
구렁텅  
시련  
격멸  
격변  
혼돈  
절망감  
번뇌도  
불실  
어둠  
허우적덜  
허우적거림  
반목

## FastText Jamo

아파올  
행진  
괴로워질  
repeat  
만드실  
gtrepeat  
섬길  
다가올  
걸어갈  
나아서  
거듭날  
아플  
overoverover  
이어질  
다스릴  
덮쳐올  
살아날  
새로워질  
neverendneverend  
끌어갈

## GloVe

고통  
두려움  
좌절  
죽음  
슬픔  
아픔  
늪  
고독  
시련  
어둠  
속  
세상  
분노  
현실  
외로움  
쾌락  
삶  
끝  
욕망  
가운데

## Swivel

좌절  
고통  
두려움  
현실  
시련  
슬픔  
실패  
아픔  
죽음  
과거  
나날  
삶  
늪  
신음  
파멸  
자신  
방황  
외침  
초라  
속

## LSA

침묵  
망상  
불신  
허영  
배경  
욕망  
싸움  
생활  
혼돈  
회색  
빛  
고통  
그늘  
파라솔  
환경  
전장  
집안  
공간  
국  
폭설  
소음

## LSA + PPMI

침묵  
망상  
허영  
불신  
배경  
욕망  
싸움  
생활  
혼돈  
회색  
빛  
고통  
그늘  
환경  
파라솔  
공간  
집안  
폭설  
전장  
국  
소음

# 쇼팽

## FastText

연습곡  
야상곡  
나단조  
마단조  
녹턴  
가장조  
다단조  
나장조  
마장조  
내림나단조  
장조  
가단조  
chopin  
소나타  
라단조  
사단조  
관현악  
바장조  
비창  
즉흥곡

## FastText Jamo

schonen  
klt  
blt  
sch  
weissen  
gefangen  
grozlige  
bltenzweige  
entfolge  
schiele  
muumliller  
frhlich  
zurucke  
maumldchen  
erweckt  
trotzdem  
verrckt  
weisse  
tiefe  
vergangen

## GloVe

왈츠  
녹턴  
야상곡  
스케르초  
연습곡  
마주르카  
발라드  
내림나단조  
촌역  
작품  
skul  
훈련병  
놀람  
나단조  
슈만  
합정역  
혜화  
op  
기호  
자작곡

## Swivel

녹턴  
chopin  
연습곡  
야상곡  
악장  
제번  
장조  
작품  
schubert  
중주  
슈베르트  
마장조  
베토벤  
첼로 모  
차르트  
내림  
다장조  
환상곡  
다단조  
보케리니

## LSA

쇼스타코비치  
연습곡  
사중주  
인벤션  
행진곡  
드보르작  
협주곡  
바단조  
가단조  
놀람  
소나타  
발트  
기호  
야상곡  
즉흥곡  
삼중주  
라단조  
나단조  
교향곡  
다단조

## LSA + PPMI

쇼스타코비치  
연습곡  
사중주  
인벤션  
행진곡  
드보르작  
협주곡  
바단조  
가단조  
놀람  
발트  
소나타  
기호  
야상곡  
즉흥곡  
삼중주  
나단조  
라단조  
교향곡  
다단조

# 인형

## FastText

목각  
인형사  
꼭두각시  
장난감  
마론  
호두  
소금인형  
상자  
보석함  
밀랍  
오스트랄로피테쿠스  
난장이  
오르골  
어리광쟁이  
유모차  
고철  
꼭두  
애송이  
배구공  
인형극

## FastText Jamo

itbut  
needhellip  
ahve  
wrongdont  
deserve  
wantill  
forbut  
done  
yould  
whatsoever  
saythat  
youdeserve  
worse  
possibly  
likemdash  
saidhellip  
mever  
utill  
butbut  
probablyve

## GloVe

꼭두각시  
밀랍  
꼬마  
목각  
사탕  
피노키오  
호두  
자켓  
요정  
어린애  
예쁜  
곰  
깨어진  
태엽  
마네킹  
조각  
형기  
모양  
인디안  
burring

## Siwivel

꼭두각시  
요정  
피노키오  
장난감  
상자  
고양이  
병정  
호두  
강아지  
꼬마  
예쁜  
여덟  
원숭이  
햄스터  
과자  
양철  
마리  
말썽쟁이  
사탕  
이빨

## LSA

태양  
악몽  
칼끝  
바늘  
탄산  
오물  
비단  
마약  
초침  
허상  
두근거림  
직선  
총알  
북극성  
힙합  
섬광  
갈증  
화음  
별빛  
장작

## LSA + PPMI

태양  
악몽  
칼끝  
바늘  
탄산  
오물  
마약  
비단  
초침  
두근거림  
허상  
직선  
총알  
북극성  
힙합  
섬광  
갈증  
별빛  
화음  
전쟁

# we

## FastText

they  
ourserlf  
uswe  
togetherlets  
togetherjust  
both  
theywe  
our  
eachothers  
togethers  
ourselves  
eachother  
wed  
togetherrise  
canwe  
whenwe  
you  
well  
weptthey  
gotwe

## FastText Jamo

they  
togetherjust  
both  
eachothers  
well  
togethers  
you  
togetherlets  
theywe  
whenwe  
uswe our  
eachother  
i  
toghether  
and  
togetherrise  
together  
togethergether  
gether  
ourselves

## GloVe

oxidize  
reaaadddy  
backhellipwhe  
n weeh  
californiahere  
can  
lifesick  
they  
famerz  
righthellipyea  
h know  
together  
wellcould  
walibomba  
could  
ll  
motorcaded  
now  
cause  
make

## Swivel

and  
were  
wed  
us  
again  
learn  
can  
never  
could  
before  
well  
all  
cause  
our  
think  
did  
then  
they  
start  
when

## LSA

they  
quotwe  
quotyou  
quotthey  
i  
accountants  
he  
cheapskates  
sailboats  
nowwe  
you  
sally  
fussin  
sentient  
cyou  
scrappin  
buti  
carolin  
laughen  
fuckboys

## LSA + PPMI

they  
quotwe  
quotyou  
quotthey  
i  
nowwe  
he  
cheapskates  
accountants  
you  
sailboats  
fussin  
sentient  
scrappin  
sally  
cyou  
buti  
carolin  
laughen  
fuckboys

# paris

## FastText

parise  
france  
parie  
rome  
luxembourg  
parish  
francetour  
mont  
loveline  
amsterdam  
londonrome  
hilton  
montel  
paradis  
frances  
parisienne  
berlin  
paradiscos  
californi  
franc

## FastText Jamo

parise  
france  
parish  
francetour  
hilton  
luxembourg  
parie  
loveline  
parissi  
montclair  
charlemagne  
franc  
johannesburg  
francine  
parisienne  
milanlouis  
kmont  
paradie  
mont  
paradis

## GloVe

senflamme  
lausanne  
veille  
tokyo  
eille  
couchent  
trimillennium  
olten  
anatole  
hilton  
girltokyo  
france  
hamburg  
mhabite  
london  
nnnnew  
rome  
leatherboys  
amsterdam  
madagascar

## Swivel

france  
berlin  
rome  
london  
beijing  
tokyo  
bordeaux  
san  
manhattan  
liverpool  
monaco  
york  
joli  
plein  
amsterdam  
italy  
rues  
plages  
belles  
au

## LSA

nashville  
jail  
atlanta  
thailand  
venice  
miami  
vegas  
amsterdam  
spain  
london  
tokyo  
chicago  
harlem  
shanghai  
cali  
richmond  
beijing  
france  
michigan  
hollywood

## LSA + PPMI

nashville  
jail  
atlanta  
thailand  
venice  
miami  
vegas  
amsterdam  
spain  
london  
chicago  
tokyo  
cali  
shanghai  
richmond  
harlem  
beijing  
michigan  
hollywood  
france



# coffee

FastText	FastText Jamo	GloVe	Swivel	LSA	LSA + PPMI
coffe	coffe	tea	tea	leather	leather
coffees	coffees	chugalugo	milk	soup	soup
coffeeshops	coffeeshops	pennyroyal	drink	whiskey	whiskey
toffee	tea	milk	cigarettes	packs	packs
coffeespoons	coffer	cakecake	liquor	rage	rage
tea	toffee	drink	cup	wealth	wealth
coffer	coffeespoons	beauolais	breakfast	hatred	pink
cofee	cocktail	dinner	sipping	cake	joints
beer	coffeebreak	drinking	ordering	joints	cake
cappuccino	milk	cup	soup	pink	hatred
coffers	chamomile	wine	spilt	ribbons	ribbons
coffeebreak	caffeine	shop	sour	codeine	codeine
breakfast	cappuccino	cake	beer	passion	passion
cappuccinos	breakfast	table	sippin	fortune	red
coffeehouses	beer	chimry	cigarette	red	fortune
wine	liquorice	bottle	stocked	labor	labor
cup	coffers	cream	lemonade	brandy	lust
caffine	soda	riededrinkin	pancakes	lemonade	grief
koffee	cofee	beer	oatmeal	grief	lemonade
milk	frappuccino	bed	smoking	lust	beer

# instagram

## FastText

pentagram  
tagram  
instagraming  
instagrams  
pentagrams  
seagram  
anagram  
instagram  
diagram  
httpinstagram  
instagrammed  
instagramming  
twitters  
timelines  
twitpic  
instagrammin  
mictwitter  
postcode  
facebooking  
demographics

## FastText Jamo

instagrams  
instagraming  
tagram  
instagrl  
instagrammed  
instargram  
httpinstagram  
instagrammin  
instagramming  
insta  
facebook  
intagram  
twitter  
facebooking  
instsgram  
sns  
pentagram  
instar  
timeline  
pic

## GloVe

thelaconicpenguin  
cmcwonderboy  
artxinb  
comdewingsofficial  
dewings  
aiueohooo  
etcdoyoung  
comjua  
commamonnscho  
httpinstagram  
comxkillerwhale  
comubenta  
httpfacebook  
facebook  
comtangerinpiano  
comptk  
comliveonchoii  
combayoberman  
twitter  
comugboz

## Swivel

youtube  
soundcloud  
facebook  
twitter  
페이스북  
httpswww  
insta  
google  
pak  
hashtag  
다운로드  
httpsyoutu  
prod  
timeline  
photo  
셀카  
www  
연예인  
검색  
트위터

## LSA

google  
whiteface  
headstone  
facebook  
astor  
rodgers  
youtube  
tumblr  
cigars  
ap  
sullivan  
louis  
jan  
kazoo  
eiderdown  
simons  
wagner  
ensemble  
sixtysix  
upholstery

## LSA + PPMI

#google  
facebook  
whiteface  
astor  
headstone  
youtube  
tumblr  
rodgers  
cigars  
ap  
sullivan  
jan  
louis  
kazoo  
eiderdown  
ensemble  
wagner  
simons  
sixtysix  
blvd

# Results

- Swivel steadily performed well regardless of languages.
- FastText(single character based) known to be suitable for Korean showed lower than expected outcomes.