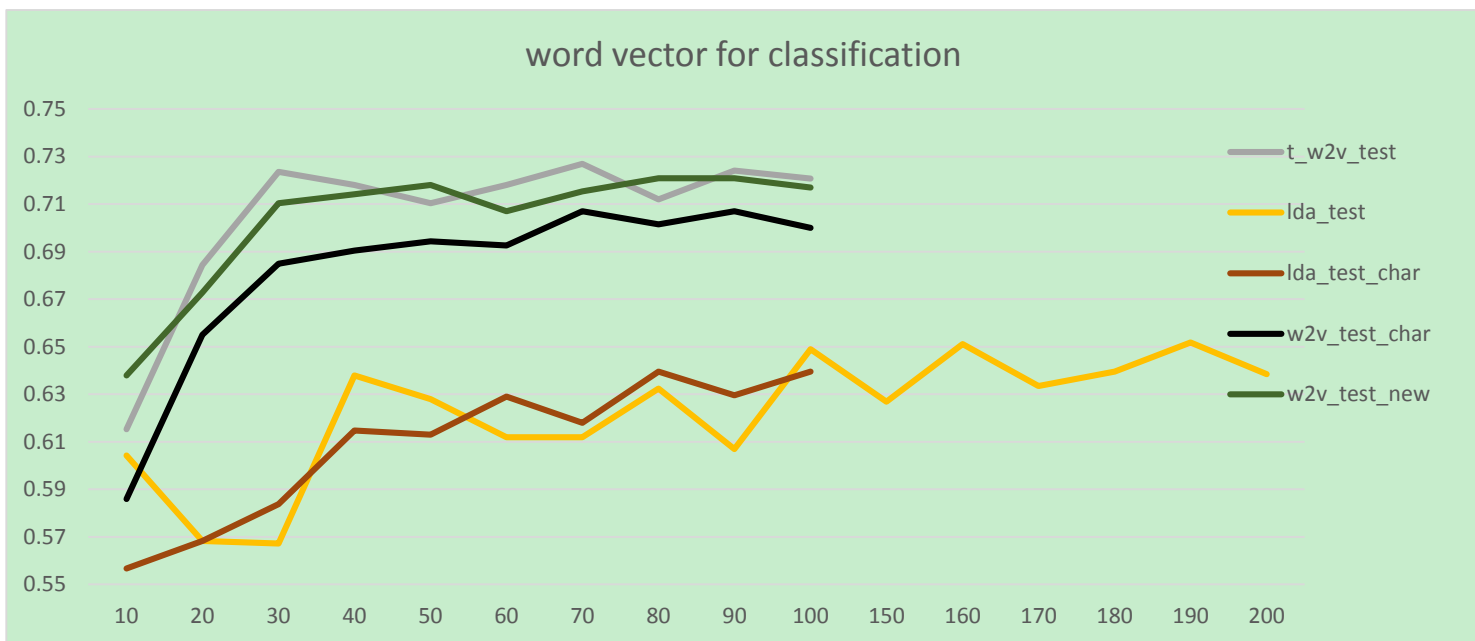


test on test-set



注：横坐标： topic num(dimension)      纵坐标： 分类的正确率

t\_w2v\_test :      train word vector on train-set(1500 documents)

w2v\_test\_char:    train word vector on renminribao-char,

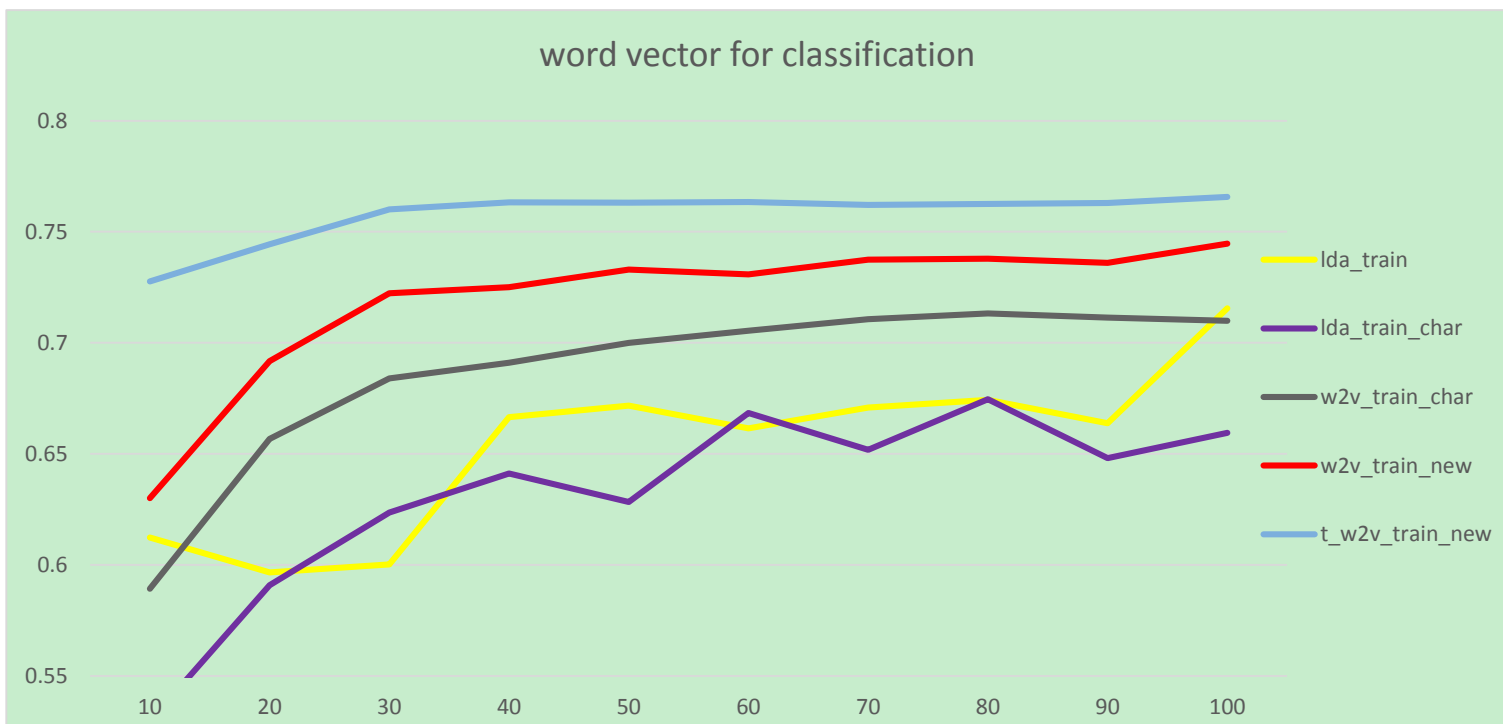
w2v\_test\_new:    train word vector on renminribao-word,

test-set:            200 documents

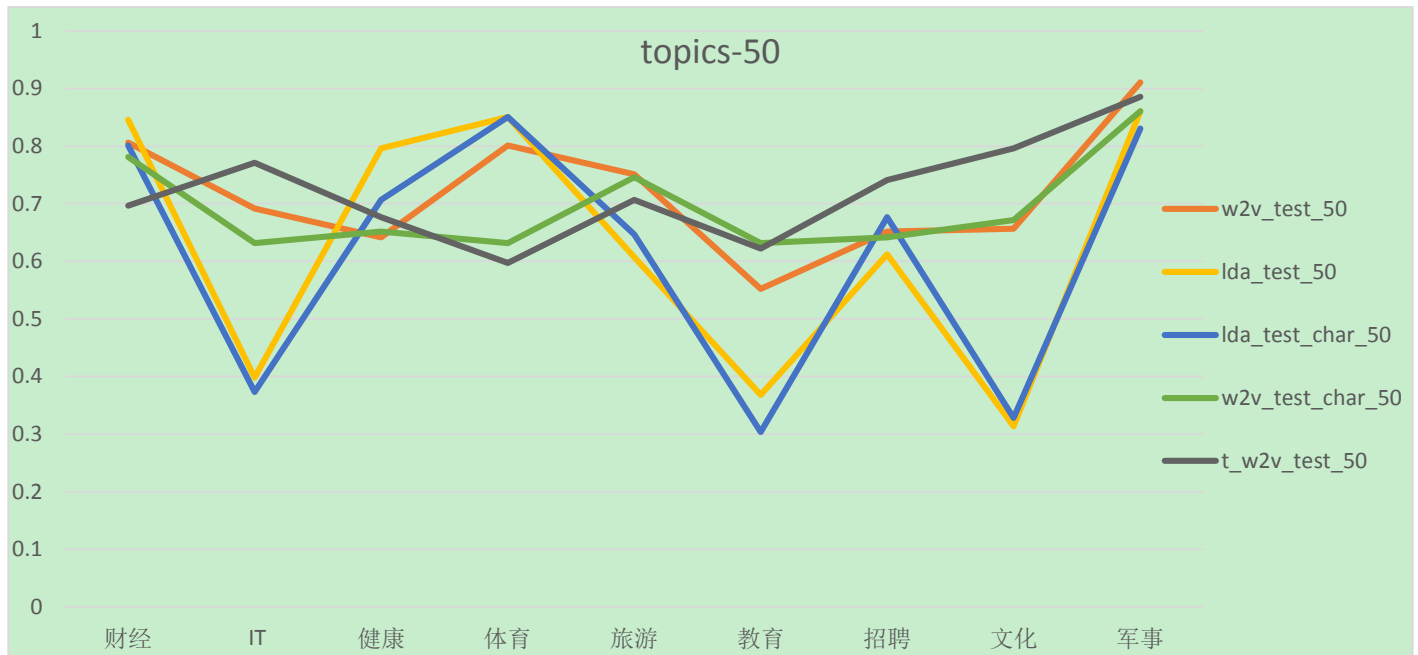
lda\_test\_char:    train topic-documents on train-set-char

lda\_test\_new:     train topic-documents on train-set-word

test on train-set



二  
test



注：横坐标：类别 纵坐标：分类的正确率

t\_w2v\_test\_50: train word vector on train-set(1500 documents)

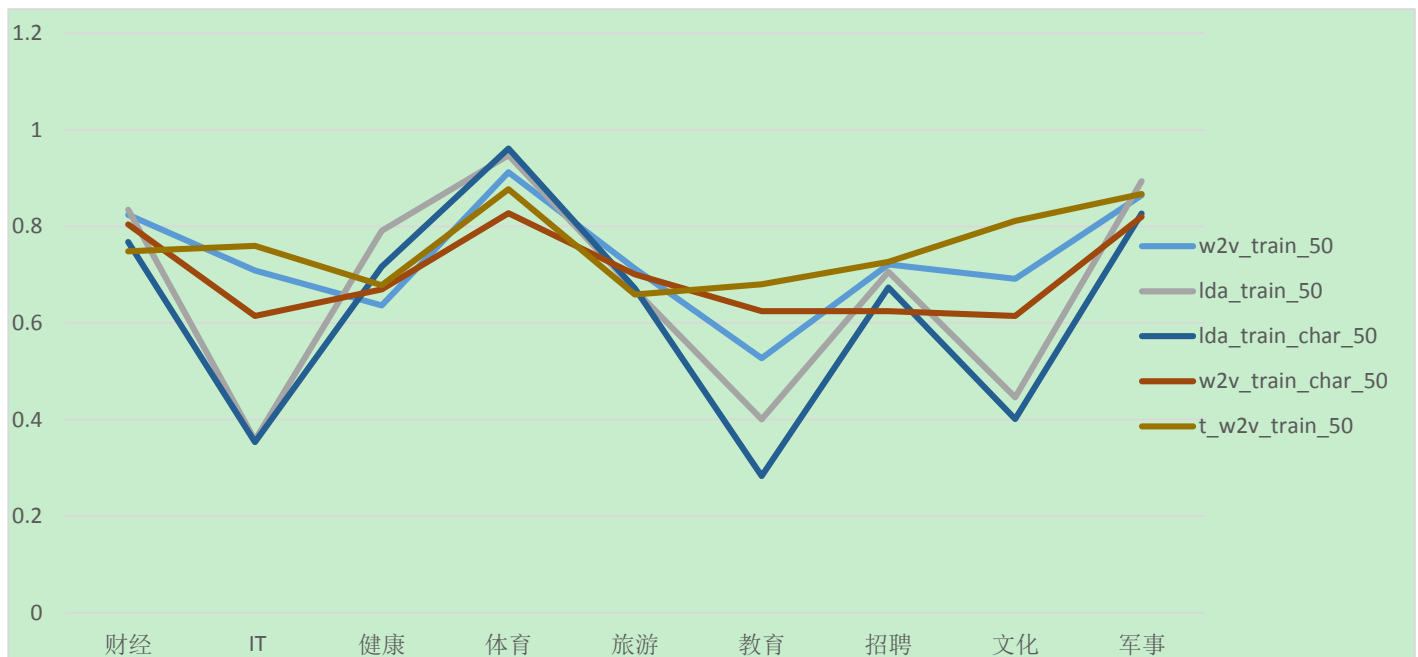
w2v\_test\_char\_50: train word vector on renminribao-char,

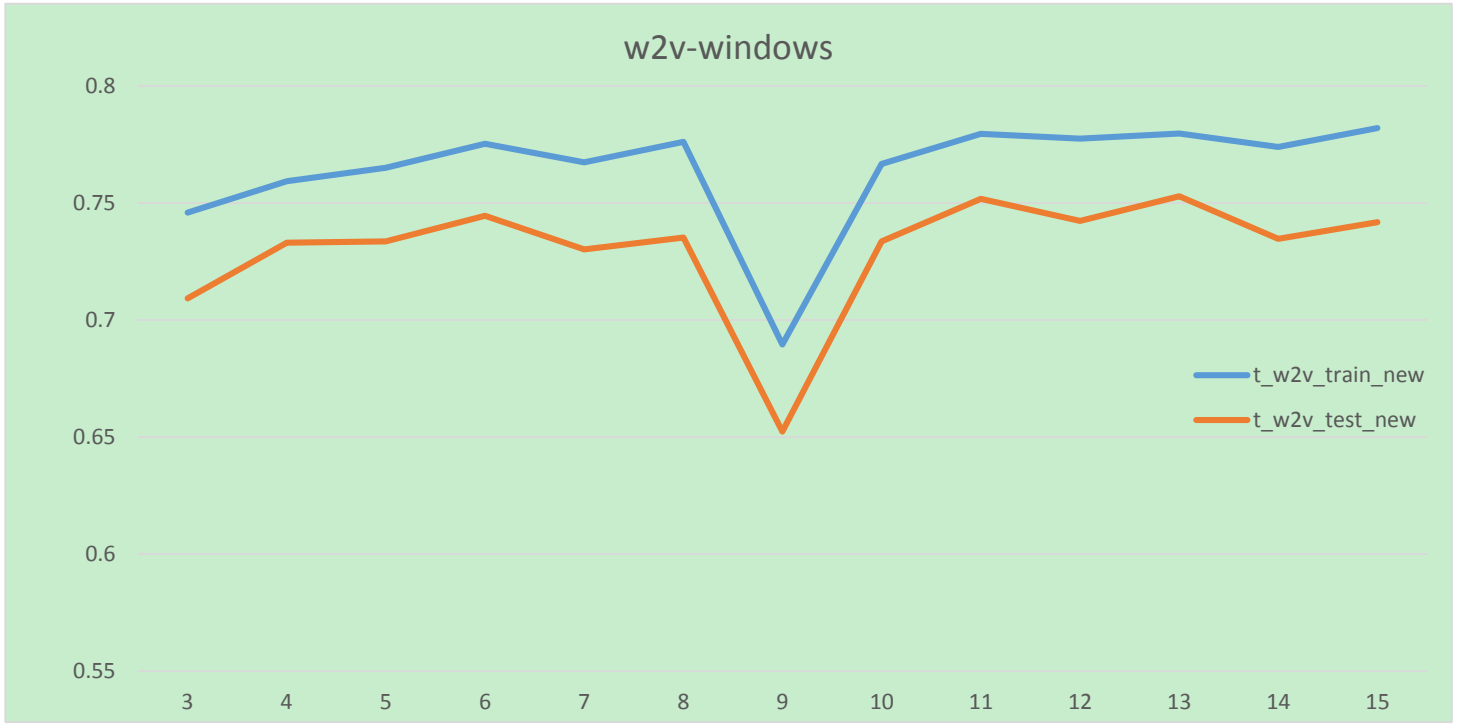
w2v\_test\_50: train word vector on renminribao-word,  
test-set: 200 documents

lda\_test\_char\_50: train topic-documents on train-set-char

lda\_test\_50: train topic-documents on train-set-word

train





注：  
横坐标： windows 大小（训练 word vector ）  
纵坐标： 分类的正确率