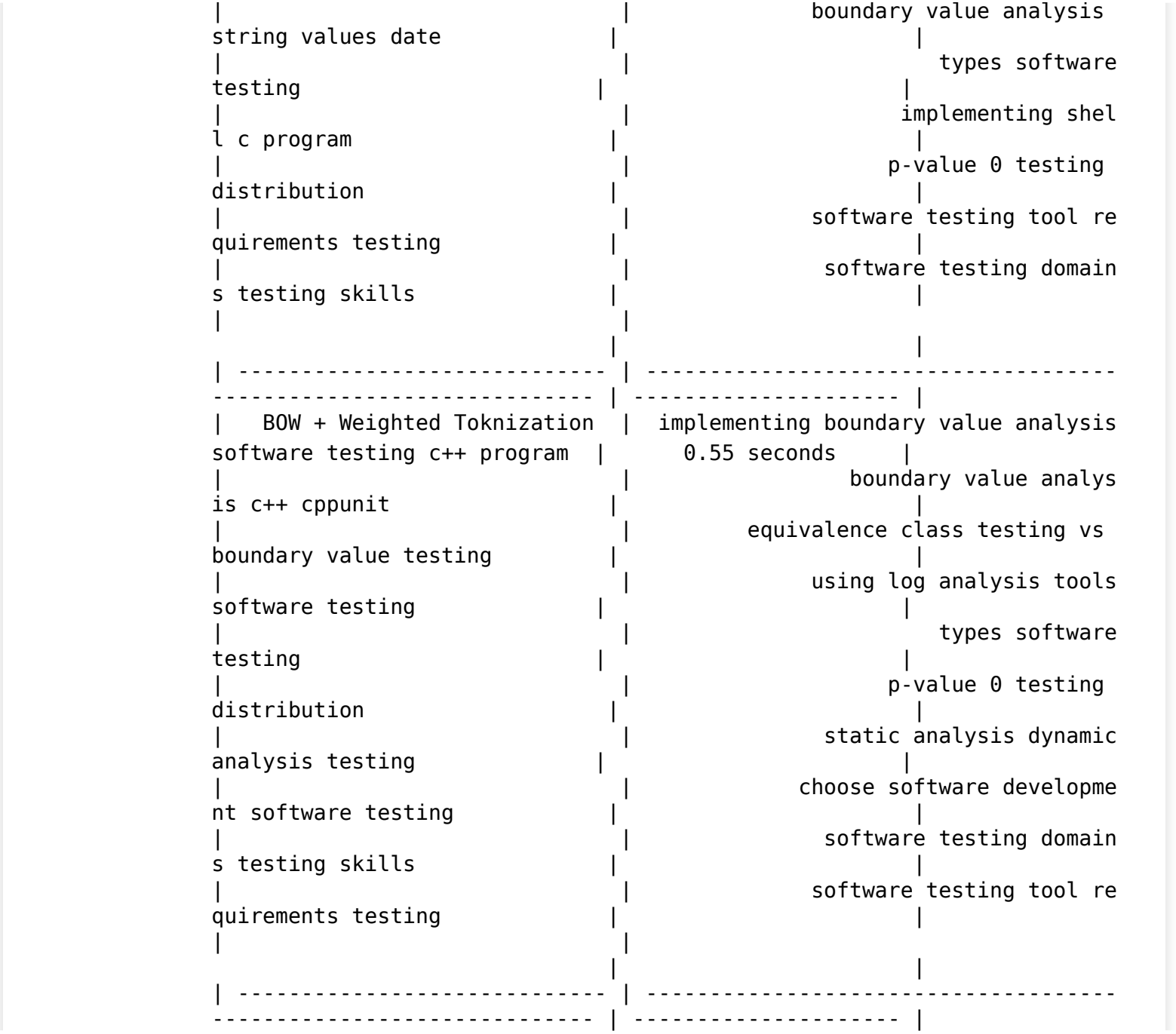
In [39]:
```
from prettytable import PrettyTable
x = PrettyTable()
x.field_names = ["Model Used", "Top 10 Results", "Query Time (Seconds)"
]
x.add_row(["BOW", "implementing boundary value analysis software testin
g c++ program\nboundary value analysis c++ cppunit\nequivalence class t
esting vs boundary value testing\nusing log analysis tools software tes
ting\nboundary value analysis string values date\ntypes software testin
g\nimplementing shell c program\np-value 0 testing distribution\nsoftwa
re testing tool requirements testing\nsoftware testing domains testing
 skills\n", '0.65 seconds'])
x.add_row(["-----------------------------", "-------------------------
----------------------------------------", '---------------------'])
x.add_row(["BOW + Weighted Toknization", "implementing boundary value a
nalysis software testing c++ program\nboundary value analysis c++ cppun
it\nequivalence class testing vs boundary value testing\nusing log anal
ysis tools software testing\ntypes software testing\np-value 0 testing
 distribution\nstatic analysis dynamic analysis testing\nchoose softwar
e development software testing\nsoftware testing domains testing skills
\nsoftware testing tool requirements testing\n", '0.55 seconds'])
x.add_row(["-----------------------------", "-------------------------
----------------------------------------", '---------------------'])
x.add_row(["TF-IDF", "implementing boundary value analysis software tes
ting c++ program\nboundary value analysis string values date\nboundary
 value analysis c++ cppunit\nusing log analysis tools software testing
\nequivalence class testing vs boundary value testing\nstatic analysis
 dynamic analysis testing\nimplementing boundary-fill algorithm opengl
\nimplementing java analysis algorithms\ntypes software testing\nimplem
enting shell c program\n", '0.41 seconds'])
x.add_row(["-----------------------------", "-------------------------
----------------------------------------", '---------------------'])
x.add_row(["TF-IDF + Weighted Toknization", "implementing boundary valu
e analysis software testing c++ program\nboundary value analysis c++ cp
punit\nusing log analysis tools software testing\nequivalence class tes
ting vs boundary value testing\nstatic analysis dynamic analysis testin
g\nboundary value analysis string values date\ntypes software testing\n
```

```
choose software development software testing\nsoftware testing tools te
sting web application\nunit testing tools generating boundary condition
s\n", '0.50 seconds'])
x.add_row(["------------------------------", "-------------------------
---------------------------------------", '--------------------'])
x.add_row(["TF-IDF-Word2Vec", "implementing boundary value analysis sof
tware testing c++ program\njustify software testing management\ncompati
bility test testing method use building software\nrest client software
 development testing\nautomated testing explaining business value\nbest
 unit testing framework testing wp7 application\ncategorise various sof
tware testing methods\nneed idea source code testing evaluation tool\nb
est practice data validation enterprise application\nextending homework
 testing platform include code analysis c c++\n", '4.2 seconds'])
x.add_row(["------------------------------", "-------------------------
---------------------------------------", '--------------------'])
x.add_row(["TF-IDF-Word2Vec + Tokens", "implementing boundary value ana
lysis software testing c++ program\njustify software testing management
\ncompatibility test testing method use building software\nrest client
 software development testing\nautomated testing explaining business va
lue\nbest unit testing framework testing wp7 application\ncategorise va
rious software testing methods\nneed idea source code testing evaluatio
n tool\nbest practice data validation enterprise application\nextending
 homework testing platform include code analysis c c++\n", '6.4 second
s'])
x.add_row(["------------------------------", "-------------------------
---------------------------------------", '--------------------'])
x.add_row(["LDA + BOW", "implementing boundary value analysis software
 testing c++ program\nexecute c++ console program c++ program\ntesting
 c++ program testing classes normally used classes\nwriting program ope
n use another program audio program\ncache hit miss value c c++ program
\nread trackbar control value c++ winapi program\nchanging value stdout
 c++ program\nadvantages c++ cli formerly managed c++ standard c++\ncom
piling small c++ program visual c++ express\nc++ program gives error va
lue initiated\n", '1.15 seconds'])
x.add_row(["------------------------------", "-------------------------
---------------------------------------", '--------------------'])
x.add_row(["LDA + TF-IDF", "implementing boundary value analysis softwa
re testing c++ program\nserendipity booksellers software program c++\nt
esting xml xmlunit variable value\nhardcode value texbox c++\nboundary
```

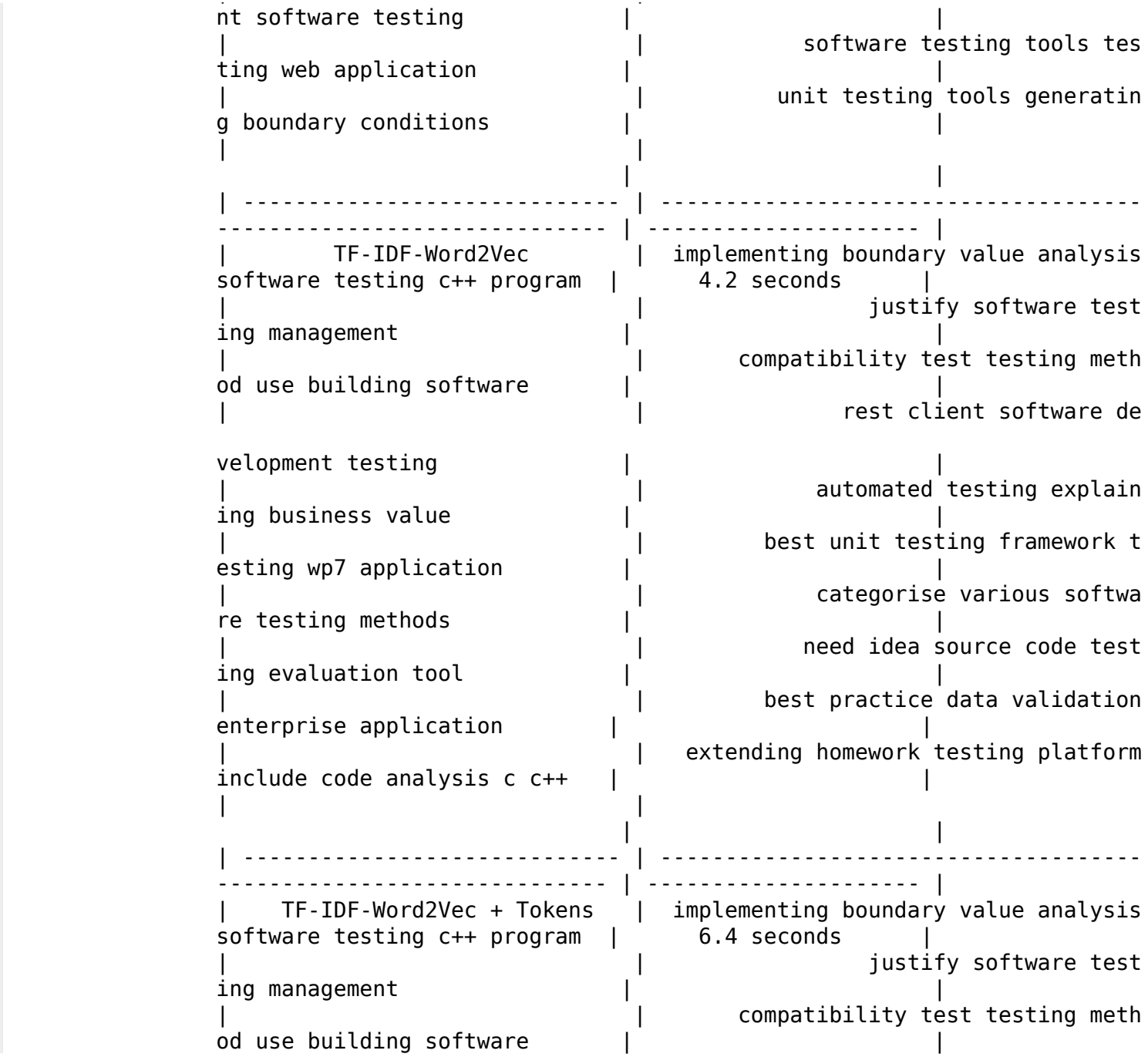
```
 value analysis c++ cppunit\nbignum divison value assignment c++\nexten
ding lifetime temporary value c++\ninsertion pair mapped value multimap
 c++\npolymorphic containers value semantics c++\nxerces c++ xml escape
 really hard", '1.39 seconds'])
x.add_row(["------------------------------", "-------------------------
---------------------------------------", '--------------------'])
x.add_row(["Universal Sentance Encoder \n+\nCosine Distance", "implemen
ting boundary value analysis software testing c++ program\nimplementing
 data structures algorithms c++\nboundary value analysis c++ cppunit\nc
ode metrics analysis unmanaged c++ code\nc++ code profiling analysis ma
c mpi\ncalculating critical path dag c++\nmethods implementing using gr
aphs nodes c++\ncomplex data structures embedding extending python c++
\nobfuscate c++ variables functions\nc++ static global non-pod theory p
ractice\n", '5.65 seconds'])
x.add_row(["------------------------------", "-------------------------
---------------------------------------", '--------------------'])
x.add_row(["Universal Sentance Encoder \n+\nEuclidean Distance", "imple
menting boundary value analysis software testing c++ program\nimplement
ing data structures algorithms c++\nboundary value analysis c++ cppunit
\ncode metrics analysis unmanaged c++ code\nc++ code profiling analysis
 mac mpi\ncalculating critical path dag c++\nmethods implementing using
 graphs nodes c++\ncomplex data structures embedding extending python c
++\nobfuscate c++ variables functions\nc++ static global non-pod theory
 practice\n", '4.87 seconds'])
print(x)
```

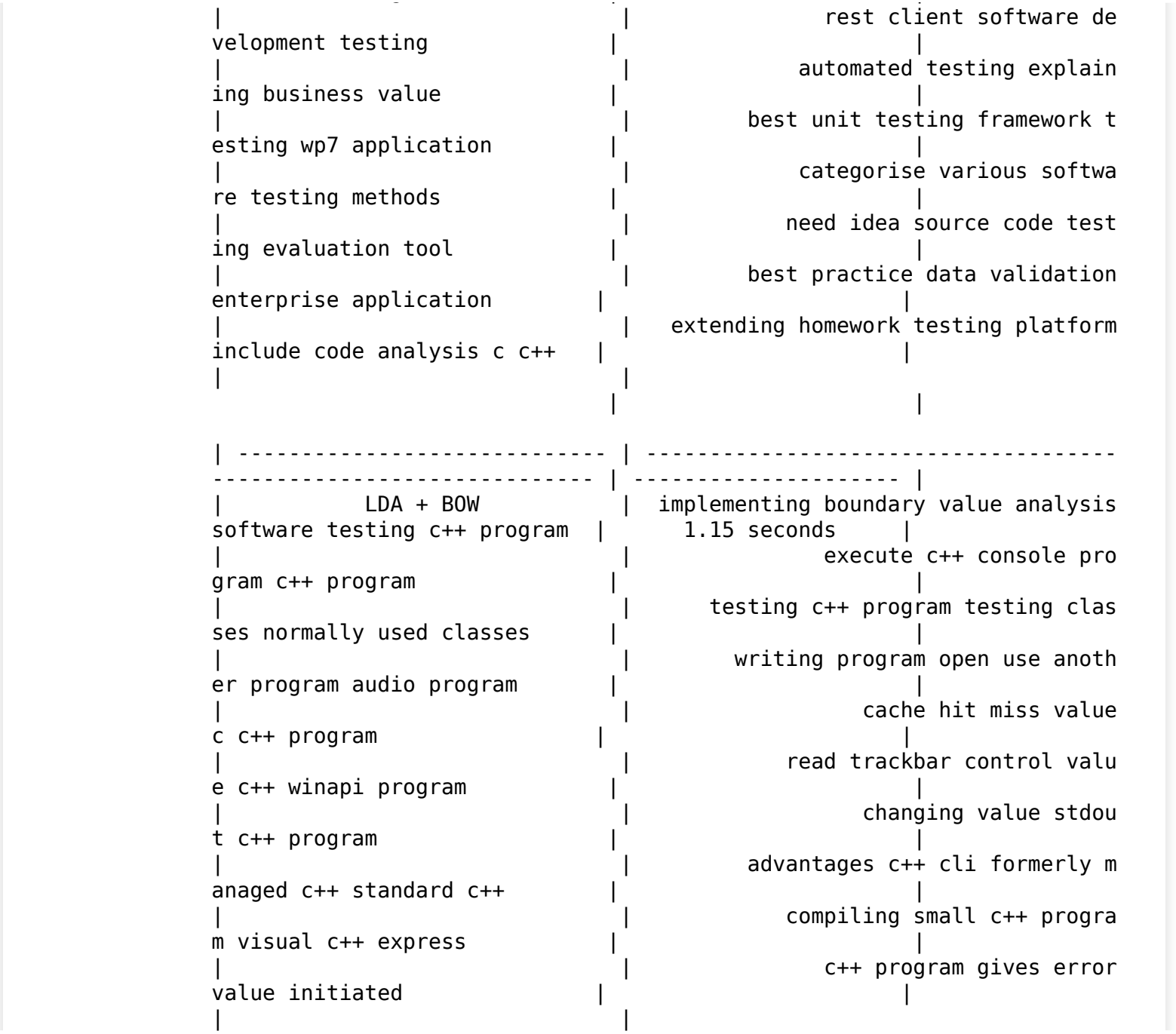
```
+------------------------------+-------------------------------------
------------------------------+----------------------+
|           Model Used         |                            Top 10 Res
ults                           |  Query Time (Seconds) |
+------------------------------+-------------------------------------
------------------------------+----------------------+
|              BOW             |  implementing boundary value analysis
software testing c++ program  |      0.65 seconds     |
|                              |                boundary value analys
is c++ cppunit                |                       |
|                              |        equivalence class testing vs
boundary value testing        |                       |
|                              |             using log analysis tools
software testing              |                       |
```

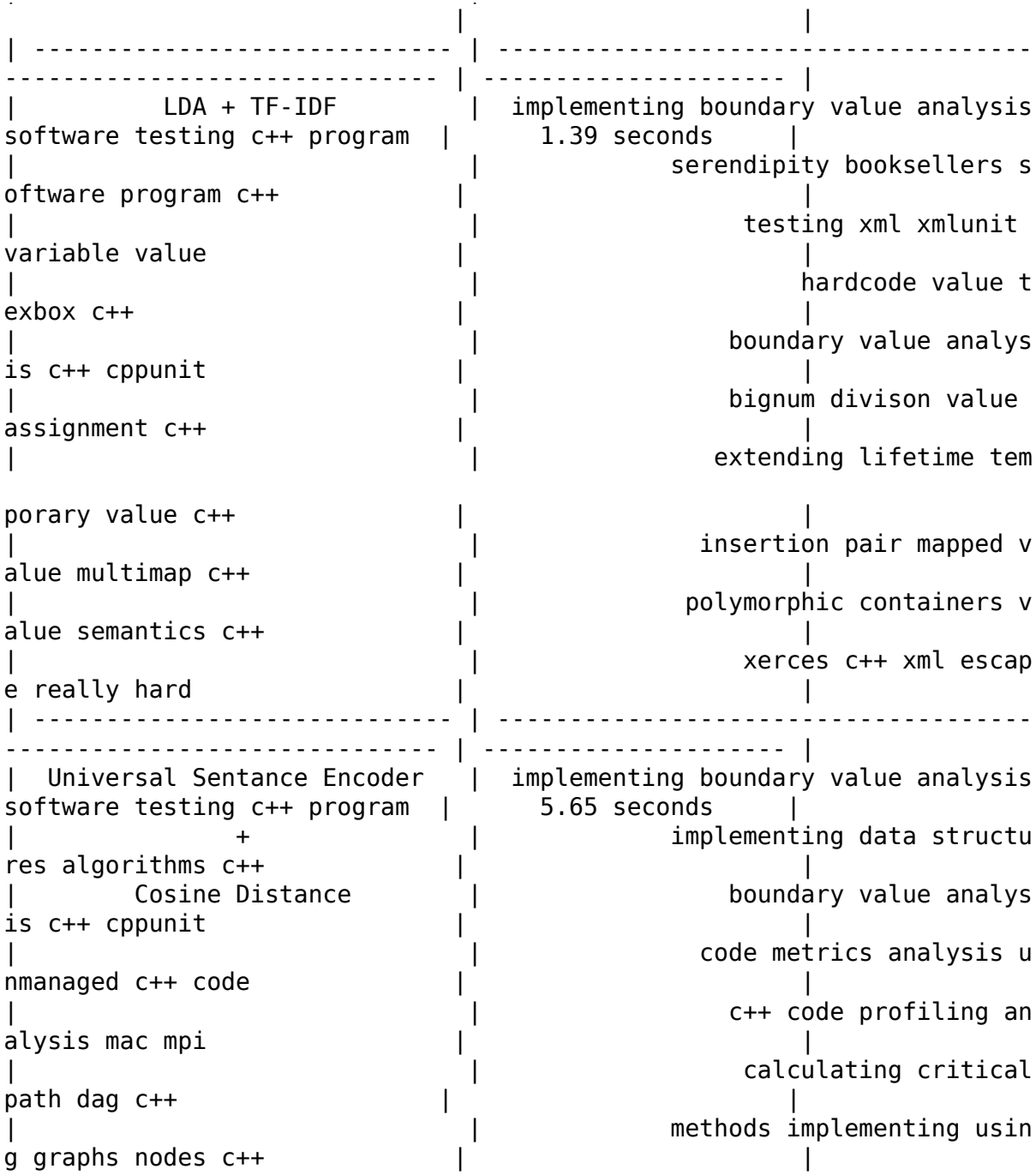


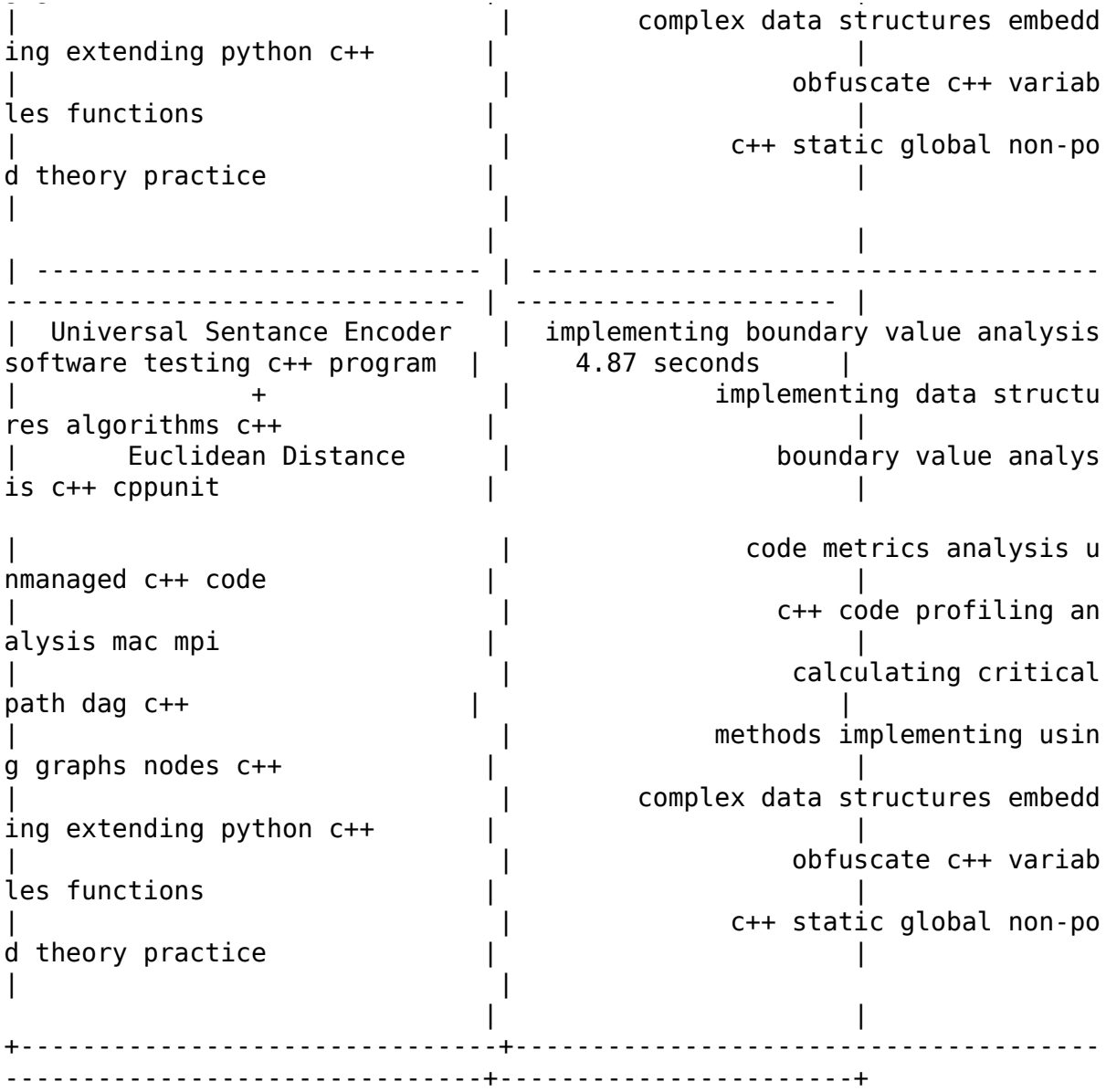
TF-IDF	implementing boundary value analysis software testing c++ program	0.41 seconds
string values date	boundary value analysis
is c++ cppunit	boundary value analys
software testing	using log analysis tools
boundary value testing	equivalence class testing vs
analysis testing	static analysis dynamic
l algorithm opengl	implementing boundary-fil
ysis algorithms	implementing java anal
testing	types software
l c program	implementing shel

TF-IDF + Weighted Toknization	implementing boundary value analysis software testing c++ program	0.50 seconds
is c++ cppunit	boundary value analys
software testing	using log analysis tools
boundary value testing	equivalence class testing vs
analysis testing	static analysis dynamic
string values date	boundary value analysis
testing	types software
	choose software developme









Summary:

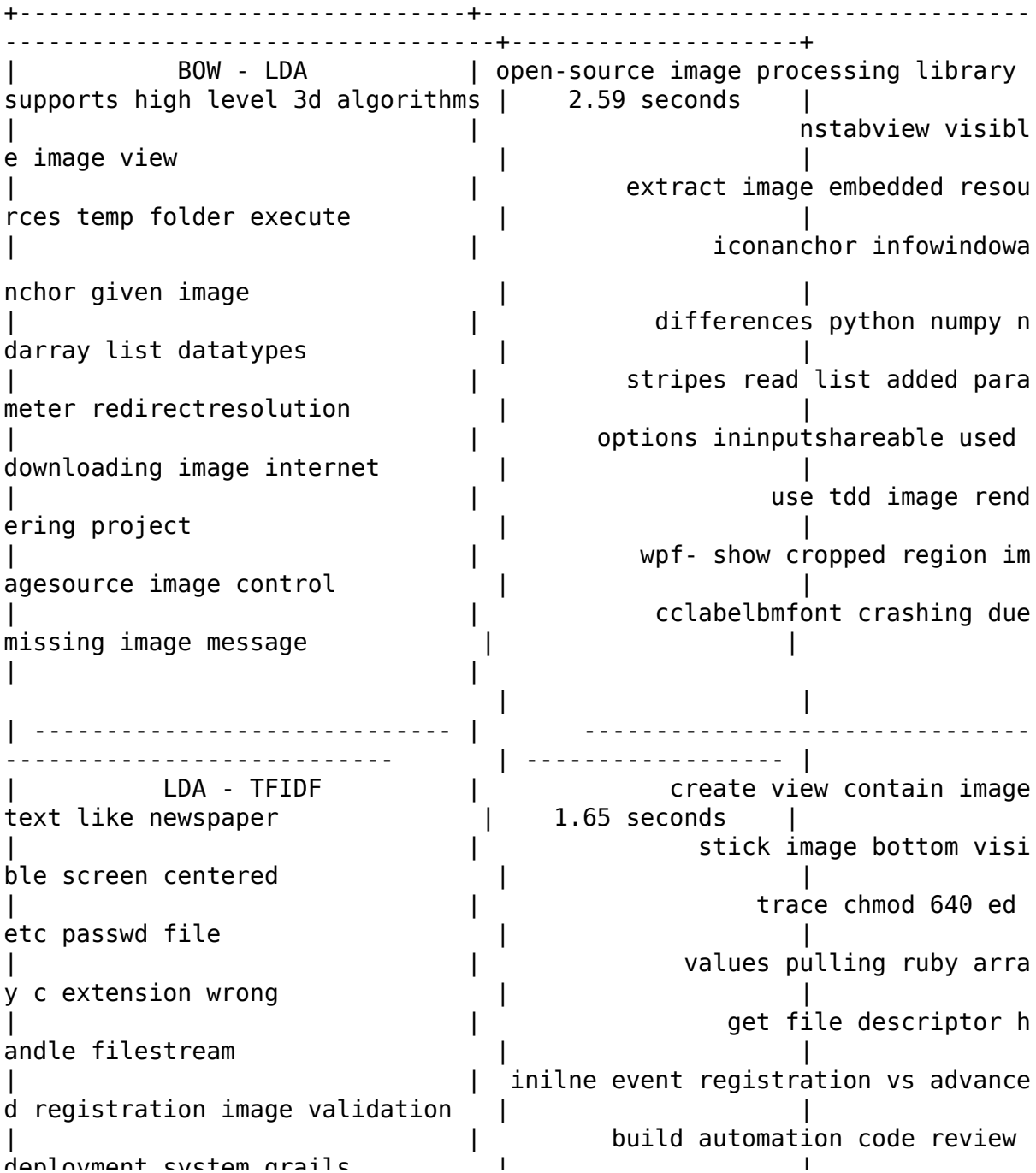
1. BOW was capable of providing us with the results that contained similar words
2. BOW + Tokenization helped us to provide results similar to BOW but the results were displayed of the same tokens or tags as we gave more weights to the tokens.
3. TF-IDF results were pretty much similar to BOW but the sequence in which the results were displayed were varied
4. TF-IDF + Tokenization provided us similar results as TF-IDF but with the same tokens or tags since they were given more weights.
5. All the results achieved were taking less time in BOW and TF-IDF but the results did not considered semantic meaning or words.
6. Hence, we tries TF-IDF Word2Vec using Glove Vectors which considered semantic meaning but the time taken to display results were bit high.
7. We tried a top modelling approach, LDA (Latent Dirichlet Allocation) using BOW and TF-IDF which considered the different topic distributions to return results.
8. Finally, we tried Universal Sentance Encoder which provided us 512 dimension embedding for each sentance. We tried Euclidean and Cosine Distance measures in this case.

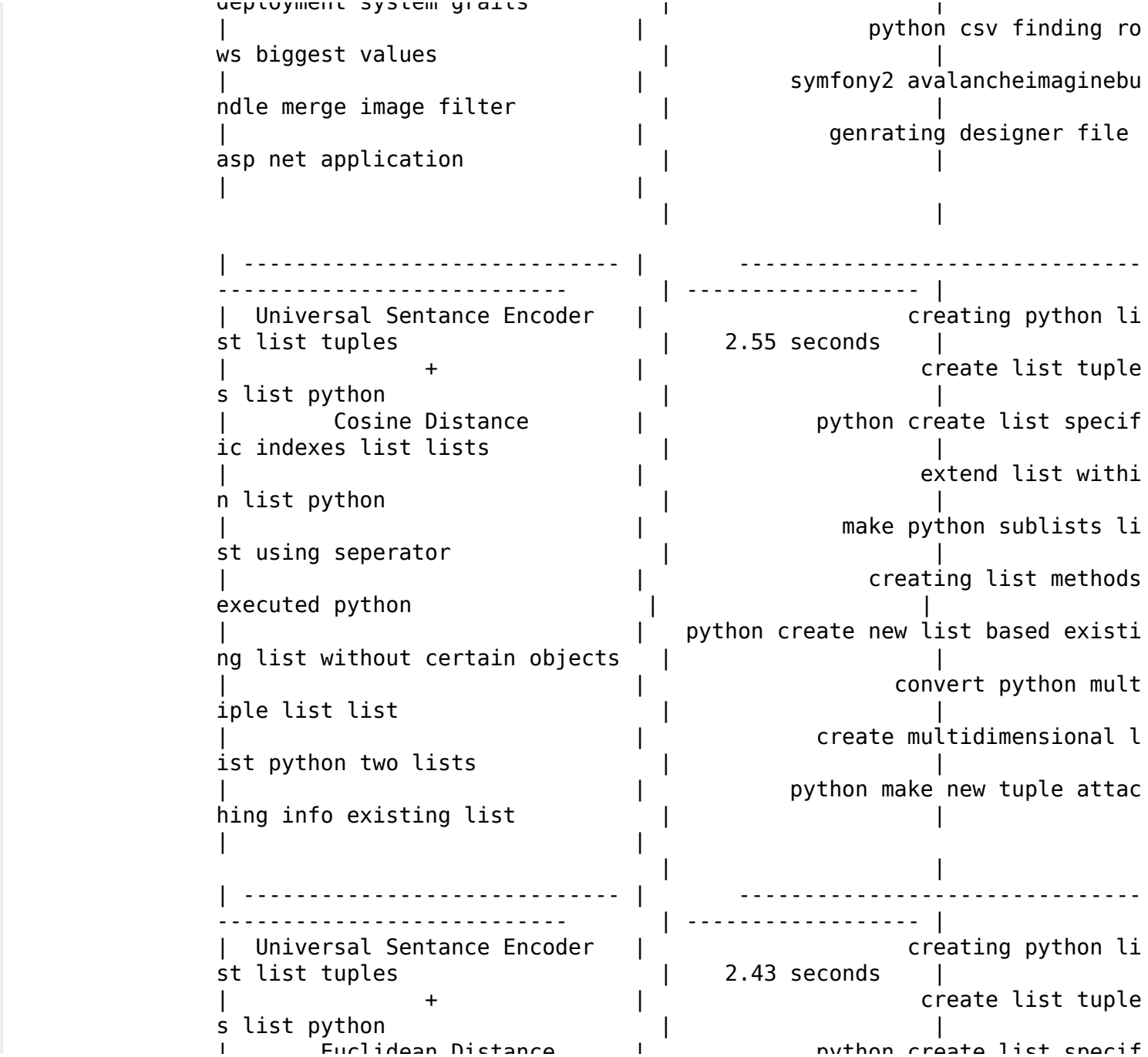
Summary for Applied AI Queries

In [24]:
```
from prettytable import PrettyTable
x = PrettyTable()
print('\nQuery Entered by user - how to create a linked list in python
\n')
x.field_names = ["Model Used", "Top 10 Results", "Query Time (Sec)"]

x.add_row(["BOW - LDA", "open-source image processing library supports
 high level 3d algorithms\nnstabview visible image view\nextract image
 embedded resources temp folder execute\niconanchor infowindowanchor gi
ven image\ndifferences python numpy ndarray list datatypes\nstripes rea
d list added parameter redirectresolution\noptions ininputshareable use
d downloading image internet\nuse tdd image rendering project\nwpf- sho
w cropped region imagesource image control\ncclabelbmfont crashing due
 missing image message\n", '2.59 seconds'])
x.add_row(["------------------------------", "-------------------------
--------------------------------", '-----------------'])
```

```
x.add_row(["LDA - TFIDF", "create view contain image text like newspape
r\nstick image bottom visible screen centered\ntrace chmod 640 ed etc p
asswd file\nvalues pulling ruby array c extension wrong\nget file descr
iptor handle filestream\ninilne event registration vs advanced registra
tion image validation\nbuild automation code review deployment system g
rails\npython csv finding rows biggest values\nsymfony2 avalancheimagin
ebundle merge image filter\ngenrating designer file asp net application
\n", '1.65 seconds'])
x.add_row(["------------------------------", "-------------------------
--------------------------------", '-----------------'])

x.add_row(["Universal Sentance Encoder \n+\nCosine Distance", "creating
 python list list tuples\ncreate list tuples list python\npython create
 list specific indexes list lists\nextend list within list python\nmake
 python sublists list using seperator\ncreating list methods executed p
ython\npython create new list based existing list without certain objec
ts\nconvert python multiple list list\ncreate multidimensional list pyt
hon two lists\npython make new tuple attaching info existing list\n",
'2.55 seconds'])
x.add_row(["------------------------------", "-------------------------
--------------------------------", '-----------------'])

x.add_row(["Universal Sentance Encoder \n+\nEuclidean Distance", "creat
ing python list list tuples\ncreate list tuples list python\npython cre
ate list specific indexes list lists\nextend list within list python\nm
ake python sublists list using seperator\ncreating list methods execute
d python\npython create new list based existing list without certain ob
jects\nconvert python multiple list list\ncreate multidimensional list
 python two lists\npython make new tuple attaching info existing list\n
", '2.43 seconds'])

print(x)
```

```
Query Entered by user - how to create a linked list in python

+------------------------------+-------------------------------------
----------------------------------+--------------------+
|           Model Used         |                                Top 10 Re
sults                            |  Query Time (Sec)  |
```





```
|       Euclidean distance      |              python create list specif
ic indexes list lists           |                     |
|                               |                   extend list withi
n list python                   |                     |
|                               |                make python sublists li
st using seperator              |                     |
|                               |                  creating list methods
executed python                 |                     |
|                               |   python create new list based existi

ng list without certain objects |                     |
|                               |                  convert python mult
iple list list                  |                     |
|                               |              create multidimensional l
ist python two lists            |                     |
|                               |           python make new tuple attac
hing info existing list         |                     |
|                               |
                                |                     |
+-------------------------------+-------------------------------------
-------------------------------+--------------------+
```

In [26]:

```
from prettytable import PrettyTable
x = PrettyTable()
print('\nQuery Entered by user - LSTM with Keras\n')
x.field_names = ["Model Used", "Top 10 Results", "Query Time (Sec)"]

x.add_row(["BOW - LDA", "disable version control features qt creator\np
roblem getting tooltip refresh properly itemrenderer flex\nerrors crawl
ing content sources sharepoint 2010\npurpose boolean switch statements
 javascript\nerrors computing psd inside parfor loops loops\nproblem ho
sting wcf service iis express\nerrors logged http https response 200\ne
rrors reported iphone mfmailcomposeviewcontroller safe\nproblem imitati
ng vs output window textbox wpf textbox\nproblem getting tomcat start r
eboot\n", '2.49 seconds'])
x.add_row(["-----------------------------", "-------------------------
--------------------------------", '-----------------'])

x.add_row(["LDA - TFIDF", "scrolling touch devices phonegap cordova pro
jects\ndelphi 2009 converts delphi 7 projects build configurations\ndeb
```

```
ug jsp pages eclipse+maven+jetty\nviewing http headers wcf service\nvie
wing contents previous commits xcode\nviewing output visual studio xslt
 debugger\ndns works iterative recursive queries\ndns service discovery
 multicast dns bonjour related\npossible varchar hash indexing structur
e mysql\nviewing uploaded pdf doc pictures without converting\n", '1.29
 seconds'])
x.add_row(["-----------------------------", "-------------------------
--------------------------------", '------------------'])

x.add_row(["Universal Sentance Encoder \n+\nCosine Distance", "2d convo
lution python similar matlab conv2\ninterpolation morphing image labvie
w opencv\ninterpolation subsampling 3d data python without vtk\nmultipl
ying matrix vector glm opengl\nopencv python bindings grabcut algorithm
\nbilinear interpolation pil image python\neasiest way perform modular
 matrix inversion python\ncalculate affine motion model coefficients us
ing opencv c++\nperform bilinear interpolation python\ncreating contour
 opencv using python\n", '2.57 seconds'])
x.add_row(["-----------------------------", "-------------------------
--------------------------------", '------------------'])

x.add_row(["Universal Sentance Encoder \n+\nEuclidean Distance", "2d co
nvolution python similar matlab conv2\ninterpolation morphing image lab
view opencv\ninterpolation subsampling 3d data python without vtk\nmult
iplying matrix vector glm opengl\nopencv python bindings grabcut algori
thm\nbilinear interpolation pil image python\neasiest way perform modul
ar matrix inversion python\ncalculate affine motion model coefficients
 using opencv c++\nperform bilinear interpolation python\ncreating cont
our opencv using python\n", '2.51 seconds'])

print(x)
```

```
Query Entered by user - LSTM with Keras

+-------------------------------+--------------------------------------
----------------------+--------------------+
|          Model Used           |                          Top 10 Results
                       |  Query Time (Sec)  |
+-------------------------------+--------------------------------------
----------------------+--------------------+
|          BOW - LDA            |         disable version control featu
```

```
res qt creator          |    2.49 seconds    |
|                              |  problem getting tooltip refresh prop
erly itemrenderer flex |                      |
|                              |       errors crawling content sources
sharepoint 2010       |                      |
|                              |         purpose boolean switch statem
ents javascript        |                      |

|                              |        errors computing psd inside pa
rfor loops loops       |                      |
|                              |           problem hosting wcf service
iis express           |                     |
|                              |            errors logged http https r
esponse 200            |                     |
|                              |   errors reported iphone mfmailcompos
eviewcontroller safe   |                      |
|                              |    problem imitating vs output window
textbox wpf textbox   |                       |
|                              |             problem getting tomcat st
art reboot             |                      |
|                              |
                       |                      |
| ---------------------------- |  ----------------------------------
---------------------- | ------------------ |
|          LDA - TFIDF         |      scrolling touch devices phonegap
cordova projects      |    1.29 seconds    |
|                              | delphi 2009 converts delphi 7 project
s build configurations |                     |
|                              |             debug jsp pages eclipse+m
aven+jetty             |                     |
|                              |                viewing http headers wc
f service              |                     |
|                              |           viewing contents previous c
ommits xcode           |                     |
|                              |          viewing output visual studio
xslt debugger         |                      |
|                              |            dns works iterative recurs
ive queries            |                     |
|                              |     dns service discovery multicast d
ns bonjour related     |                     |
|                              |        possible varchar hash indexing
```

```
|                                |        possible varchar nash indexing
structure mysql       |                    |
|                                |     viewing uploaded pdf doc pictures
without converting    |                    |
|                                |
                         |                    |
| ---------------------------- |  ----------------------------------

---------------------- | ------------------ |
|  Universal Sentance Encoder   |          2d convolution python simila
r matlab conv2         |    2.57 seconds    |
|                +               |         interpolation morphing image
labview opencv         |                    |
|        Cosine Distance        |     interpolation subsampling 3d data
python without vtk    |                    |
|                                |            multiplying matrix vector
glm opengl             |                    |
|                                |           opencv python bindings grab
cut algorithm          |                    |
|                                |           bilinear interpolation pil
image python           |                    |
|                                |     easiest way perform modular matri
x inversion python     |                    |
|                                | calculate affine motion model coeffic
ients using opencv c++ |                    |
|                                |            perform bilinear interpola
tion python            |                    |
|                                |            creating contour opencv u
sing python            |                    |
|                                |
                         |                    |
| ---------------------------- |  ----------------------------------
---------------------- | ------------------ |
|  Universal Sentance Encoder   |          2d convolution python simila
r matlab conv2         |    2.51 seconds    |
|                +               |         interpolation morphing image
labview opencv         |                    |
|       Euclidean Distance      |     interpolation subsampling 3d data
python without vtk    |                    |
|                                |            multiplying matrix vector
```

```
glm opengl                |                    |
|                         |          opencv python bindings grab
cut algorithm             |                    |
|                         |          bilinear interpolation pil
image python              |                    |
|                         |     easiest way perform modular matri
x inversion python        |                    |

|                         | calculate affine motion model coeffic
ients using opencv c++    |                    |
|                         |           perform bilinear interpola
tion python               |                    |
|                         |             creating contour opencv u
sing python               |                    |
|                         |
                          |                    |
+-------------------------+--------------------------------------
-----------------------+--------------------+
```

Summary:

From the above two queries experimented on four different models, we see that the universal sentance encoder is giving more relevant results. Hence, we consider Universal Sentance Encoder as our final model.