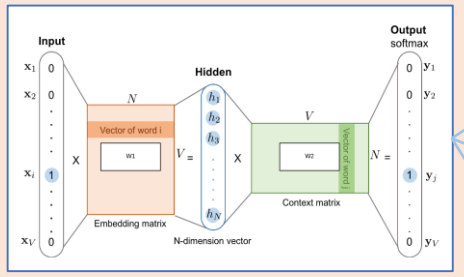
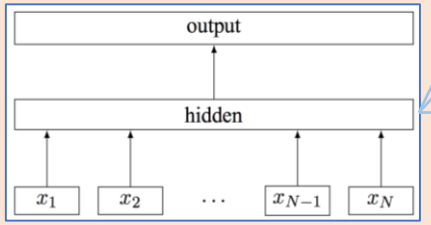
Computational Analysis: Phase 1

1 Choose Continuous Space Word Representation Models

Google Word2Vec



Facebook FastText



2 Choose Multi-Domain Corpora

- ukWaC
- English Gigaword v.5
- WaCkypedia

3 Train each model on each corpus

- ukWaC x Word2Vec
- Gigaword x Word2Vec
- WaCkypedia x Word2Vec
- ukWaC x fastText
- Gigaword x fastText
- WaCkypedia x fastText

Internally Acquired Food Norms (200)

2	almond
3	anchovy
4	apple
5	apricot
6	artichoke
7	asparagus
8	avocado
9	bacon
10	bagel
11	bamboo_shoot
12	banana
13	barley
14	basil
15	bean
16	beef
17	beet
18	berry
19	black_bean
20	blackberry
21	blueberry
22	bread
23	bread_stick
24	broccoli
25	brown_bread
26	brussel_sprout
27	butter
28	cabbage
29	capon
30	carrot
31	catfish
32	cauliflower
33	caviar
34	cereal
35	cheese
36	cherry

- Word2Vec x ukWaC (100d)
- Word2Vec x English Gigaword (100d)
- Word2Vec x WaCkypedia (100d)
- fastText x ukWaC (100d)
- fastText x English Gigaword (100d)
- fastText x WaCkypedia (100d)

100-dimensional word embeddings per model per corpus

4 Extract 200 food item vectors from vectorized corpus

Computational Analysis: Phase 2

Multi-linear Regression + 10Fold Cross Validation

Independent variable *x:*
Food item embedding

```
[array([ 0.453611, -0.366176, -0.141542, -0.893425,  0.465599,  0.392229,
        -0.507111, -0.694245, -0.032502, -0.07994 , -0.070681, -0.27096 ,
        -0.003226,  0.288651,  0.328331, -0.336236, -0.042413, -0.301928,
        -0.163257, -0.094638,  0.462432,  0.3566  ,  0.045621, -0.085892,
         0.467205,  0.45132 ,  0.348264, -0.314366,  0.195946, -0.099086,
        -0.179906, -0.539623, -0.383221, -0.075982,  0.068214, -0.741099,
        -0.34751 , -0.479687, -0.030655, -0.186941,  0.705656,  0.51782 ,
        -0.178176,  0.255177,  0.436854,  0.08742 , -0.193607, -0.075158,
        -0.686341,  0.257652,  0.19651 , -0.096867, -0.173365,  0.420043,
        -0.095786, -0.200198,  0.137489, -0.558253, -0.813881, -0.199163,
         0.266628, -0.588016,  0.049463,  0.001984,  0.510038, -0.323774,
         0.773042, -0.568593, -0.026827,  0.23736 ,  0.010912, -0.010211,
         0.681059, -0.131046, -0.311557,  1.119529,  0.195949, -0.733487,
        -0.63452 , -0.628856,  0.141703, -0.219126,  0.334232, -0.325517,
        -0.117375, -0.952969,  0.145928, -0.190898, -0.179893, -0.268953,
        -0.455921,  0.387328,  0.266881, -0.252641,  0.242846, -0.068556,
        -0.506825,  0.168405, -0.346399,  0.150842]),
```

x 200 100d vectors

$$y = m + b_1 x_1 + \cdots + b_{100} x_{100}$$

Dependent variable *y:*
participant scores

```
array([3.42105263, 4.95238095, 1.54166667, 3.85      , 5.15384615,
       5.55      , 5.1875    , 2.28571429, 3.        , 3.71428571,
       1.8       , 2.46666667, 3.35      , 2.22222222, 4.30434783,
       2.75      , 2.86363636, 2.6       , 2.52380952, 3.55555556,
       2.54545455, 2.        , 2.125     , 2.56521739, 2.81818182,
       2.68181818, 2.19047619, 5.16666667, 1.52380952, 6.75      ,
       2.22727273, 8.57142857, 1.63636364, 3.71428571, 3.86363636,
       2.7826087 , 2.05555556, 4.625     , 3.        , 3.04545455,
       3.625     , 7.1875    , 5.21428571, 3.72      , 4.15      ,
       3.61111111, 2.66666667, 3.3125    , 1.6       , 2.23529412,
       3.95238095, 6.5       , 3.31818182, 2.31578947, 2.13043478,
       3.47058824, 1.73913044, 6.83333333, 2.66666667, 5.8       ,
       4.54545455, 1.59090909, 4.27272727, 7.75      , 4.2       ,
       4.6       , 5.15789474, 4.94444444, 3.35      , 7.16666667,
       3.61904762, 3.5       , 4.53846154, 5.23809524, 6.625     ,
       2.52380952, 3.1       , 2.45454545, 2.42105263, 4.54545455,
       2.94444444, 5.        , 2.5       , 3.26315789, 4.55555556,
       3.80952381, 5.15384615, 3.30769231, 2.22222222, 3.0952381 ,
       5.        , 4.66666667, 4.77272727, 2.34782609, 2.52380952,
       6.35714286, 2.54545455, 1.625     , 3.        , 6.33333333,
       7.94736842, 3.36842105, 3.        , 4.        , 1.83333333,
       2.29411765, 6.4       , 2.66666667, 2.66666667, 6.15384615,
       2.9       , 3.        , 1.47619048, 1.63157895, 7.33333333,
       4.65217391, 1.86363636, 1.92      , 2.19230769, 4.75      ,
       2.57894737, 2.6       , 6.        , 2.41666667, 1.47619048,
       2.14285714, 2.52380952, 2.72727273, 3.60869565, 7.25      ,
       5.        , 3.44      , 3.64285714, 3.5625    , 2.91304348,
       3.84210526, 4.42857143, 2.94444444, 1.38095238, 3.5       ,
       2.63157895, 5.68181818, 1.90909091, 6.        , 2.5       ,
       2.4375    , 3.18181818, 8.        , 4.25      , 6.        ,
       3.125     , 2.55      , 2.31578947, 3.09756098, 3.        ,
       1.52380952, 4.11764706, 4.86363636, 5.81818182, 2.375     ,
       2.35294118, 2.63636364, 6.94117647, 5.125     , 3.875     ,
       3.875     , 6.2       , 3.35294118, 3.61111111, 2.4       ,
       2.71428571, 2.        , 6.25      , 6.        , 4.08695652,
       1.9047619 , 2.15      , 7.4       , 4.4375    , 1.375     ,
       1.80952381, 6.91666667, 5.55      , 8.5625    , 2.9047619 ,
       7.125     , 3.86363636, 6.75      , 7.58823529, 2.08333333,
       3.40909091, 2.91666667, 2.10526316, 1.875     , 3.58823529,
       1.66666667, 3.        ])
```

Result: *y_predicted:*

```
[ 4.31269781  5.00135427  3.44086883  1.40455785  4.06720328  5.81442979
  4.08207388  2.11727744  5.3071059   5.58519795  3.4475461   0.57784826
  4.4902382   2.00288621  3.42518456  3.17412859  3.38084598  2.91337976
  2.37247592  3.1551803   2.52560599  2.76121753  1.77403474  4.33484236
  4.9515539   2.6445249   2.06136429  6.81265225  1.24924578  7.55271351
  2.16348738  6.73845265  2.41125085  2.51321669  3.03336963  4.30665572
  2.02774951  5.14953959  3.24031312  3.56200914  2.91268389  2.98276473
  3.29858793  3.51164121  2.68555889  5.37707204  6.03084055  3.75151635
  1.26309016  2.08025484  3.94004556  5.32562254  0.82024393  4.46426824
  2.09941095  4.46093842  0.69652597  3.10867791  3.93000656  3.9420461
  5.35694452  3.34872474  4.40609539  6.08251001  2.30409775  2.85067007
  5.79897798  5.060091    4.19030632  5.54774665  0.21380597 -0.0398384
  5.73000945  6.10661685  5.61566724  6.16410365  3.16412241  2.96041644
  3.96139041  4.33853659  2.83062592  4.46047913  3.14478753  4.34133971
  4.21541991  3.78499832  7.83429438  2.2940397   1.32810556  4.20233464
  2.26749432  5.67350059  3.89830654  1.34737306  4.24806122  6.42102869
  1.51413048  3.05222362  3.05175646  4.79902155  5.39394163  2.95306756
  4.8551654   3.29465991  2.51592146  4.04850793  6.89642733  1.2864659
  4.05802261  4.51090791  4.18186638  3.40445846  2.92527036  0.85902943
  7.50547622  3.0027805   1.65466891  2.03765615  0.96674214  3.49129612
  4.64279707  0.69180296  6.44288857  2.56635572  1.79133345  2.97752764
  3.00776242  2.16142534  1.36212465  5.98757662  3.26167682  3.47388633
  3.59718123  4.73936994  3.55221345  4.190833    4.9133889   4.4231573
  1.02534345  5.29843424  0.92750393  5.03136374  1.70037422  6.08729535
  0.59711136  2.11439328  1.82487475  7.31956746 10.92227686  4.48324249
  5.33242691  0.78368467  1.46171959  3.05855933  5.0981172   1.70556363
  2.90233067  3.3132541   6.47470758  2.22689537  2.61760758  3.81672579
  9.09726041  5.36124404  2.34866855  2.84289089  3.76741552  2.49699906
 -0.66984611  1.4311968   2.228245    3.13283515  5.67011167  4.49531783
  2.70190559  1.57126584  2.04603086  4.20929992  3.59152727  3.53608431
  1.7181082   5.67475537  5.78658435  7.61843043  5.5890636   7.55895155
  4.01135369  5.61438433  7.76125018  3.3774434   4.76924465  3.1935754
  2.61686604  2.49687677  5.83687034  3.63833084  3.95799437]
```