Michael Covarrubias
PID#: A12409694
Naveen Ketagoda
PID#: A10773459

Homework 5

Problem 1

1.1

- We are able to determine for any value of $c > 0$ $K'(x,z) = cK(x,z)$ is a kernel given $K(x,z)$ is a kernel. Proving that a kernel multiplied by a positive scalar is still a kernel. Where every component of the feature map is just multiplied (or scaled) by c. Therefore this IS a kernel

1.2

- We did not determine a feature map for $c < 0$ $K'(x,z) = cK(x,z)$ because even though $K(x,z)$ is a kernel, a negative v value would cause $K'(x,z)$ to violate the PSD requirement for the kernel matrix to meet it it were to be considered a kernel. Therefore this is NOT a kernel

1.3

- Using the solution from previous problem (1.1) we conclude that both kernels being multiplied by positive values result in valid kernels. Now the question is if the summation between two kernels is also a kernel. To validate this we concluded the feature map of the two summed kernels. In the supplementary work we determine that we can conclude the feature map to be the inner product between
- $<[(\sqrt{c1})\boldsymbol{\Phi}1(x) \;\; (\sqrt{c2})\boldsymbol{\Phi}2(x)], [(\sqrt{c1})\boldsymbol{\Phi}1(z) \;\; (\sqrt{c2})\boldsymbol{\Phi}2(z)]>$. Therefore this IS a kernel

1.4

- With K1 and K2 both being kernels, the questions on hand now is determining if the product of two kernels is also a kernel. Both K1 and K2 being kernels means that they both satisfy returning a scalar. And the product of a scalar with another scalar is also a scalar, so with this logic intuition would have us believe that $K'(x,z)$ is also a kernel. Through the supplementary work we concluded that the feature map for $K'(x,z)$ = would be $<[\boldsymbol{\Phi}1(x)\boldsymbol{\Phi}2(x)], [\boldsymbol{\Phi}1(z)\boldsymbol{\Phi}2(z)]>$ where the feature map for $\boldsymbol{\Phi}'(x) = [\boldsymbol{\Phi}1(x)\boldsymbol{\Phi}2(x)]$. Therefore this IS a kernel

Problem 2

2.1

- For problem 2.1 we first check to see if $K(x,z) = x1z2$ meets the requirements of being a kernel before we continue in attempt to determine the feature map for $K(x,z)$. First check is to check for symmetry for all x and z where $K(x,z) = K(z,x)$. In the example included in our homework we have shown a case where $K(x,z) \neq K(z,x)$ having is conclude that $K(x,z) = x1z2$ is NOT a kernel

2.2

- For problem 2.2, as in most cases, we start by checking if $K(x,z)$ meets both requirements of symmetry and PSD before we go on to determine the feature map.

Looking at K(x,z) = 1 - <x,z> it does satisfy the symmetry requirement. Where K(x,z) = 1 - <x,z> = 1 - <z,x> = K(z,x). We now move on to checking the PSD requirement. In the included example of our homework we show the example for x being a 1x1 scalar not equal to 0 which would cause t'Kt to be less than 0 where t is also a 1x1 scalar not equal to 0.

2.3

- Problem 2.3 we are not bounded by what type of function F is. Therefore the only conclusion we can say is that in some cases F can be a kernel, and in other cases f is not a kernel. In our homework we give both examples of F being a kernel and also F not being a kernel where it would violate either symmetry or PSD. But for All cases of what F can be K(x,z) = F(x1,x2)F(z1,z2) is NOT a kernel

2.4

- For the proof of 2.4 we recall back to the lecture where we had a problem set where we proved K(x,z) = min(x,z) to be a kernel a useful correlation. We also refer back to Problem 1.3 where we concluded that the summation of two kernels is also a kernel to be helpful. Because K(x,z)= Σ min(xi,zi) to just be a summation of *d* kernels as we have shown in the feature map determined in our homework. Therefore K(x,z)= Σ min(xi,zi) IS a kernel

2.5

- For this problem we approached it by determining the feature map for an example where x and z are both 3x1 vectors. Through this example we were able to find a corresponding term for each possible choice of subset which helped us prove that K(x, z) = (1 + x1z1)(1 + x2z2) . . . (1 + xdzd) IS a kernel because we can determine the feature map for x and z being dx1 vectors

2.6

- For problem 2.6 we refer back to the problem set we had in class where we determined K'(x,z) = max(x,z) to NOT be a kernel. A deeper look at 2.6 we can break it down to it being a summation of *d* number of Non kernels. But to prove this we continued to determine if K(x,z)= Σ max(xi,zi) meets the symmetry and PSD requirement in order to be considered a kernel. Our example in our homework shows that K(x,z)= Σ min(xi,zi) violates PSD. concluding in K(x,z)= Σ min(xi,zi) NOT being a kernel

2.7

- In order to solve 2.7 we referred back to the proof in the lecture notes of the string kernel. Where in the example in the notes is over an alphabet of letters and in our case it's a dictionary of words using this correlation in the notes we are able to determine the $\boldsymbol{\Phi}$(x) to have a coordinate for all possible worlds in the dictionary D. If word u is a word in string x then $\boldsymbol{\Phi}$u(x) = 1 and $\boldsymbol{\Phi}$u(x) = 0 otherwise. Our homework ran through the complete feature map $\boldsymbol{\Phi}$(x) where $\boldsymbol{\Phi}$(x) includes a coordinate for every possible word in the dictionary D.

Problem 3

3.1

- Listed below are the error’s calculated over the training and test data with p = 2, 3, 4, & 5 for string kernel Kp(s, t). Kp(s, t) is the number of substrings of length p that are common to both s and t, where a string that occurs *a* times in s and *b* times in t is counted as *a*b*
-

	Train Error	Test Error
P = 2	0.071349862259	0.0817941952507
P = 3	0.0123966942149	0.0408970976253
P = 4	0.0068870523416	0.0263852242744
P = 5	0.0068870523416	0.0343007915567

3.2

- Listed below are the error’s calculated over the training data and test data with p = 2, 3, 4, & 5 for string kernel Mp(s,t). Mp(s,t) is the number of substrings of length p that are common to both s and t, where a string that occurs *a* times in s and *b* times in t is counted as 1
-

	Train Error	Test Error
P = 2	0.0818181818182	0.0963060686016
P = 3	0.0126721763085	0.0540897097625
P = 4	0.00743801652893	0.0290237467018
P = 5	0.0068870523416	0.0343007915567

3.3

- We ran out of time and could not finish problem 3.3. However we were able to determine the W_t for p = 5 where W_t is an array of the the indices of the x_i and y_i values of the training data
- W_t = 0, 1, 2, 5, 6, 8, 13, 15, 16, 17, 18, 19, 20, 22, 24, 25, 26, 27, 28, 29, 30, 33, 34, 36, 37, 39, 42, 47, 49, 51, 53, 54, 55, 60, 62, 66, 68, 74, 75, 76, 79, 80, 81, 83, 85, 86, 88, 89, 90, 93, 98, 101, 103, 106, 109, 111, 113, 115, 116, 118, 119, 122, 123, 127, 136, 139, 143, 151, 153, 154, 155, 157, 162, 163, 164, 165, 167, 173, 175, 176, 178, 180, 187, 197, 207, 209, 213, 218, 219, 222, 231, 232, 241, 245, 254, 258, 259, 261, 263, 265, 272, 280, 281, 282, 284, 292, 297, 304, 310, 316, 318, 326, 329, 339, 344, 345, 358, 360, 369, 377, 378, 391, 394, 396, 397, 405, 411, 413, 414, 421, 423, 426, 427, 432, 438, 439, 441, 448, 457, 461, 462, 468, 470, 476, 478, 489, 490, 500, 501, 505,

523, 526, 534, 549, 554, 560, 569, 572, 597, 600, 611, 613, 620, 638, 641, 646, 648, 660, 670, 671, 675, 697, 708, 713, 715, 723, 733, 748, 761, 770, 777, 781, 783, 785, 791, 794, 804, 813, 841, 846, 855, 856, 858, 859, 872, 887, 895, 938, 951, 962, 987, 1000, 1001, 1016, 1028, 1031, 1038, 1039, 1046, 1052, 1055, 1056, 1057, 1066, 1067, 1072, 1084, 1089, 1092, 1095, 1097, 1099, 1105, 1113, 1116, 1121, 1139, 1141, 1148, 1151, 1153, 1154, 1188, 1192, 1197, 1201, 1221, 1222, 1226, 1231, 1237, 1243, 1244, 1263, 1278, 1294, 1297, 1299, 1307, 1308, 1337, 1344, 1355, 1356, 1364, 1371, 1373, 1376, 1390, 1391, 1397, 1402, 1404, 1407, 1410, 1418, 1422, 1426, 1429, 1438, 1441, 1443, 1444, 1448, 1459, 1473, 1483, 1485, 1489, 1497, 1499, 1519, 1526, 1527, 1537, 1539, 1550, 1552, 1570, 1574, 1597, 1605, 1611, 1625, 1635, 1642, 1649, 1663, 1679, 1693, 1703, 1712, 1731, 1732, 1738, 1742, 1744, 1748, 1760, 1761, 1762, 1767, 1781, 1789, 1792, 1805, 1819, 1830, 1842, 1853, 1854, 1858, 1874, 1876, 1895, 1900, 1906, 1929, 1941, 1945, 1948, 1954, 1955, 1975, 1991, 2022, 2061, 2063, 2077, 2078, 2086, 2089, 2116, 2120, 2128, 2136, 2143, 2188, 2196, 2198, 2199, 2207, 2249, 2273, 2300, 2321, 2325, 2328, 2350, 2352, 2359, 2363, 2368, 2372, 2373, 2375, 2377, 2385, 2387, 2389, 2390, 2393, 2403, 2422, 2445, 2447, 2452, 2456, 2457, 2459, 2489, 2508, 2515, 2531, 2548, 2571, 2576, 2603, 2606, 2613, 2627, 2629, 2637, 2664, 2686, 2725, 2732, 2735, 2753, 2756, 2783, 2785, 2815, 2821, 2823, 2826, 2831, 2857, 2875, 2906, 2909, 2915, 2928, 2941, 2953, 2959, 2966, 2967, 2970, 3087, 3089, 3095, 3101, 3104, 3134, 3145, 3147, 3150, 3156, 3162, 3163, 3171, 3189, 3199, 3204, 3207, 3208, 3211, 3219, 3222, 3227, 3232, 3244, 3254, 3257, 3263, 3267, 3269, 3283, 3284, 3309, 3322, 3326, 3343, 3344, 3359, 3365, 3368, 3391, 3410, 3418, 3419, 3434, 3435, 3441, 3465, 3468, 3472, 3494, 3512, 3555, 3572, 3623, 3625

- **The following results below are my output for problem 3.3 where I have 11 strings tied for the 1st largest value and 22 values ties for the 2nd largest value. The following solutions were determined after the deadline for partial credit as I talked to Professor Chaudhuri today after class where she told me I would still be able to receive partial credit for Problem 3.3**

- 1st Maximum Values
- ['KPYKC', 'KCPEC', 'PECGK', 'PYKCP', 'YKCPE', 'CGKSF', 'HQRTH', 'GKSFS', 'CPECG', 'ECGKS', 'EKPYK']
- 2ns Maximum Values
- ['GGITW', 'DGGIT', 'RVKEV', 'QKHQR', 'QRTHT', 'RTHTG', 'TGEKP', 'SFSQS', 'GEKPY', 'HTGEK', 'THTGE', 'KSFSQ', 'FSQSS', 'KHQRT', 'LQKHQ', 'RDDTA', 'DTALN', 'DDTAL', 'ILTGL', 'GKEGH', 'CGKEG', 'SAXXS']