Haonan Hu

3/25/2021

2863545

If we label our sets of features:   
A = {z1, SL, SW, PL, PW}  
B = {z1, z2, SW, PL, PW}  
C = {z1, z2, z3, SW, PL}   
D = {z1, z2, z3, z4, SW}   
E = {z1, z2, z3, z4,SL}   
S = { z1, z2, z3, z4, SL, SW, PL, PW}

* 1. What is A ∪ E?

**{z1, SL, SW, PL, PW, z2, z3, z4}**

* 1. What is D ∩ E?

**{z1, z2, z3, z4}**

* 1. Create a mutation of C by adding another feature.

**C = {z1, z2, z3, SW, PL}, s = z4, C’ = C, s = {z1, z2, z3, z4, SW, PL}**

* 1. Create a mutation of C by deleting one feature.

**C = {z1, z2, z3, SW, PL}, C’ = C – {PL} = {z1, z2, z3, SW}**

* 1. Create a mutation of C by replacing one feature with another.

**C = {z1, z2, z3, SW, PL}, s = {z4}, C’ = C – {PL} U s = { z1, z2, z3, SW, z4}**

* 1. Assuming each k-fold cross-validation takes 0.1 sec and the rest of the GA processing time is negligible, how long will it take to evaluate one generation?

**5 selected sets for new generation, 20 cross over sets from 5 new generation sets, 5 mutations of the ones selected for new generation, 20 mutations of the new generation cross over sets, with total of 50 sets to eval under k-fold, time = 50 \* 0.1 = 5s**

* 1. With the same assumption from 6, but if we start with 6 sets of features and select the 6 best for each new generation, how long will it take to evaluate one generation?

**6 selected sets for new generation, 30 cross over sets from 6 new generation sets, 6 mutations of the ones selected for new generation, 30 mutations of the new generation cross over sets, with total of 72 sets to eval under k-fold, time = 72 \* 0.1 = 7.2s**