

# NCTU Pattern Recognition, Homework 4

0886004 周芝妤

## Part. 1, Coding (80%):

In this coding assignment, you need to implement the cross-validation and grid search by using only NumPy, then train the [SVM model from scikit-learn](#) on the provided dataset and test the performance with testing data. Find the sample code and data on the GitHub page [https://github.com/NCTU-VRDL/CS\\_ILE5065/tree/main/HW4](https://github.com/NCTU-VRDL/CS_ILE5065/tree/main/HW4)

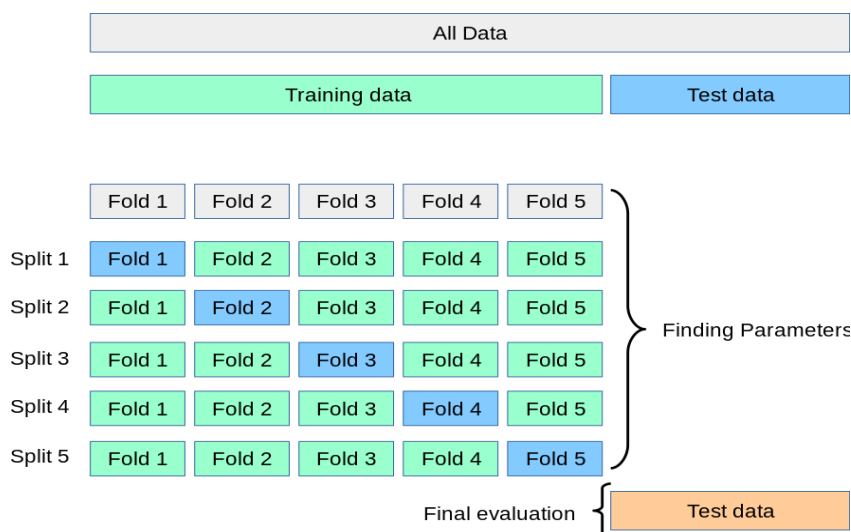
Please note that only NumPy can be used to implement cross-validation and grid search. You will get no points by simply calling [`sklearn.model\_selection.GridSearchCV`](#).

- (10%) K-fold data partition: Implement the K-fold cross-validation function. Your function should take K as an argument and return a list of lists (*len(list) should equal to K*), which contains K elements. Each element is a list contains two parts, the first part contains the index of all training folds (index\_x\_train, index\_y\_train), e.g. Fold 2 to Fold 5 in split 1. The second part contains the index of validation fold, e.g. Fold 1 in split 1 (index\_x\_val, index\_y\_val)

Note: You need to handle if the sample size is not divisible by K. Using the strategy from [sklearn](#). The first  $n\_samples \% n\_splits$  folds have size  $n\_samples // n\_splits + 1$ , other folds have size  $n\_samples // n\_splits$ , where  $n\_samples$  is the number of samples,  $n\_splits$  is K,  $\%$  stands for modulus,  $//$  stands for integer division. See this [post](#) for more details

Note: Each of the samples should be used **exactly once** as the validation data

Note: Please shuffle your data before partition



Ans: 10-Fold index is showed following.

Split: 1, Training index: [121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352 290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60 275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450 439

541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273 20  
411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142 501  
73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434 358  
11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423 244  
147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43 242  
234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492 537  
158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387 438  
167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511 107  
267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29 403  
258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219 328  
293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355 375  
169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4 235  
497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418 413  
190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279 34  
227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388 325  
311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406 225  
182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396 393  
316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359372]

Split: 2, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273 20  
411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142 501  
73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434 358  
11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423 244  
147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43 242  
234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492 537  
158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387 438  
167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511 107  
267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29 403  
258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219 328  
293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355 375  
169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4 235  
497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418 413

190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279 34  
227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388 325  
311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60 275  
516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450 439 541]

Split: 3, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423 244  
147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43 242  
234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492 537  
158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387 438  
167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511 107  
267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29 403  
258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219 328  
293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355 375  
169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4 235  
497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418 413  
190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279 34  
227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388 325  
311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273 20  
411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142 501 73  
140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434 358 11 249]

Split: 4, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 377 533 391 373 18 278 473 286 170 437 216 301 214 387 438  
167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511 107  
267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29 403  
258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219 328  
293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355 375  
169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4 235  
497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418 413  
190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279 34  
227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388 325  
311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [452 24 305 255 25 367 466 351 55 499 326 317 246 171 423 244 147  
65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43 242 234 318  
189 456 464 521 433 362 211 339 493 89 479 374 407 193 492 537 158 106 549]

Split: 5, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423

244 147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43  
242 234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492  
537 158 106 549 343 144 435 319 96 59 207 241 110 347 507 32 219 328  
293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355 375  
169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4 235  
497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418 413  
190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279 34  
227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388 325  
311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [377 533 391 373 18 278 473 286 170 437 216 301 214 387 438 167 37  
28 472 222 531 327 500 425 30 10 306 315 260 33 360 511 107 267 330 345  
196 307 38 174 135 155 266 455 431 80 414 118 496 29 403 258 400 224 518]

Split: 6, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423  
244 147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43  
242 234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492  
537 158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387  
438 167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511  
107 267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29  
403 258 400 224 518 109 1 467 449 83 180 5 178 194 408 262 418 413  
190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279 34  
227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388 325  
311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282

226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [343 144 435 319 96 59 207 241 110 347 507 32 219 328 293 13 19  
137 277 200 58 274 547 310 302 353 136 148 166 517 355 375 169 416 138 237  
64 42 76 253 488 181 101 308 153 71 261 459 4 235 497 47 185 300 67]

Split: 7, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423  
244 147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43  
242 234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492  
537 158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387  
438 167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511  
107 267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29  
403 258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219  
328 293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355  
375 169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4  
235 497 47 185 300 67 508 528 105 361 443 534 102 469 265 491 184 283  
120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363 344  
176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281 215  
444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [109 1 467 449 83 180 5 178 194 408 262 418 413 190 231 98 490  
9 39 390 503 329 514 428 149 513 82 548 379 279 34 227 365 162 86 151  
509 495 332 348 54 247 15 502 108 161 417 388 325 311 191 195 323 532 78]

Split: 8, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60

275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423  
244 147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43  
242 234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492  
537 158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387  
438 167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511  
107 267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29  
403 258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219  
328 293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355  
375 169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4  
235 497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418  
413 190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279  
34 227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388  
325 311 191 195 323 532 78 378 334 520 480 366 156 6 271 371 152 475  
48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527 282  
226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422 294  
198 124 504 131 259 160 203 17 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [508 528 105 361 443 534 102 469 265 491 184 283 120 97 157 357 41  
478 272 230 256 192 485 447 544 526 26 91 363 344 176 409 524 8 208 23  
232 368 280 238 457 44 543 252 296 201 281 215 444 133 288 115 122 519 99]

Split: 9, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406  
225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423  
244 147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43  
242 234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492  
537 158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387  
438 167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511  
107 267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29  
403 258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219  
328 293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355  
375 169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4

235 497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418  
413 190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279  
34 227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388  
325 311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184  
283 120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363  
344 176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281  
215 444 133 288 115 122 519 99 62 114 529 481 486 218 52 50 84 85  
460 150 239 432 471 45 453 245 289 380 540 483 461 127 159 95 356 376  
129 309 333 104 81 204 512 313 482 448 197 331 87 530 93 470 295 398  
370 364 236 446 338 69 56 369 221],

Validation index: [378 334 520 480 366 156 6 271 371 152 475 48 546 536 112 53 61  
177 100 442 187 538 382 395 350 248 304 527 282 226 468 487 206 94 250 240  
346 451 14 436 354 291 484 341 298 422 294 198 124 504 131 259 160 203 17]

Split: 10, Training index: [427 125 321 312 205 340 335 498 539 505 223 412 202 123 406

225 182 243 401 213 68 146 410 63 420 525 72 35 394 49 51 210 128 297 396  
393 316 130 79 324 134 419 506 16 175 88 426 117 132 285 383 116 462 359  
372 121 233 46 251 276 22 154 111 337 477 188 314 494 77 143 90 352  
290 269 342 299 287 209 36 229 74 199 119 421 445 264 173 145 165 60  
275 516 284 322 404 3 320 257 523 179 2 75 402 228 489 126 542 450  
439 541 217 424 212 57 183 392 66 465 458 384 292 397 220 40 441 273  
20 411 389 405 515 522 268 12 254 349 139 440 463 429 7 70 474 142  
501 73 140 545 399 430 164 385 163 386 476 510 381 92 141 186 113 434  
358 11 249 452 24 305 255 25 367 466 351 55 499 326 317 246 171 423  
244 147 65 0 303 27 21 31 263 415 454 535 168 172 336 270 103 43  
242 234 318 189 456 464 521 433 362 211 339 493 89 479 374 407 193 492  
537 158 106 549 377 533 391 373 18 278 473 286 170 437 216 301 214 387  
438 167 37 28 472 222 531 327 500 425 30 10 306 315 260 33 360 511  
107 267 330 345 196 307 38 174 135 155 266 455 431 80 414 118 496 29  
403 258 400 224 518 343 144 435 319 96 59 207 241 110 347 507 32 219  
328 293 13 19 137 277 200 58 274 547 310 302 353 136 148 166 517 355  
375 169 416 138 237 64 42 76 253 488 181 101 308 153 71 261 459 4  
235 497 47 185 300 67 109 1 467 449 83 180 5 178 194 408 262 418  
413 190 231 98 490 9 39 390 503 329 514 428 149 513 82 548 379 279  
34 227 365 162 86 151 509 495 332 348 54 247 15 502 108 161 417 388  
325 311 191 195 323 532 78 508 528 105 361 443 534 102 469 265 491 184  
283 120 97 157 357 41 478 272 230 256 192 485 447 544 526 26 91 363  
344 176 409 524 8 208 23 232 368 280 238 457 44 543 252 296 201 281  
215 444 133 288 115 122 519 99 378 334 520 480 366 156 6 271 371 152  
475 48 546 536 112 53 61 177 100 442 187 538 382 395 350 248 304 527  
282 226 468 487 206 94 250 240 346 451 14 436 354 291 484 341 298 422  
294 198 124 504 131 259 160 203 17],

Validation index: [ 62 114 529 481 486 218 52 50 84 85 460 150 239 432 471 45 453  
245 289 380 540 483 461 127 159 95 356 376 129 309 333 104 81 204 512 313  
482 448 197 331 87 530 93 470 295 398 370 364 236 446 338 69 56 369 221]



2. (30%) Grid Search & Cross-validation: using [sklearn.svm.SVC](#) to train a classifier on the provided train set and conduct the grid search of “C” and “gamma”, “kernel”=’rbf’ to find the best hyperparameters by cross-validation. Print the best hyperparameters you found.

Note: We suggest use K=5

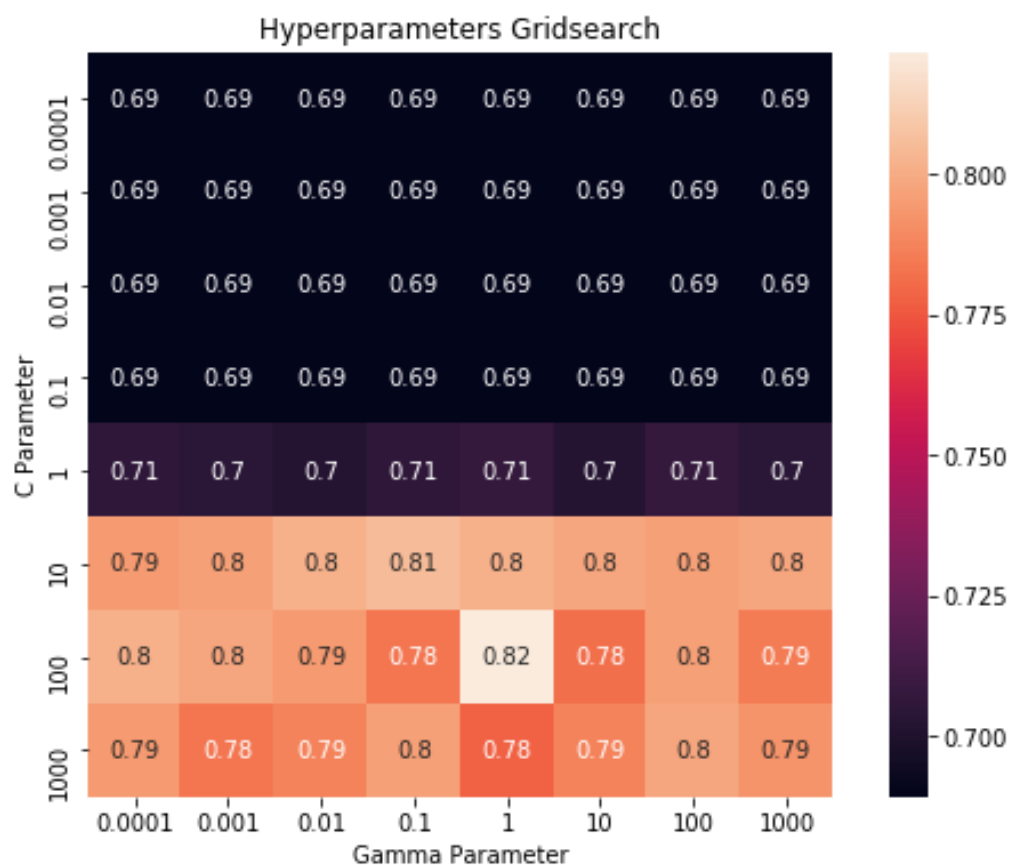
Ans: K=5, Best score is 0.81 in training, Best parameters: {'gamma': 0.001, 'C': 100}

3. (10%) Plot the grid search results of your SVM. The x, y represent the hyperparameters of “gamma” and “C”, respectively. And the color represents the average score of validation folds.

Note: This image is for reference, not the answer

Note: [matplotlib](#) is allowed to use

Ans: I also use seaborn to plot this heatmap.



4. (15%) Train your SVM model by the best hyperparameters you found from question 2 on the whole training set and evaluate the performance on the test set.

Note: Your accuracy scores should be higher than 0.85

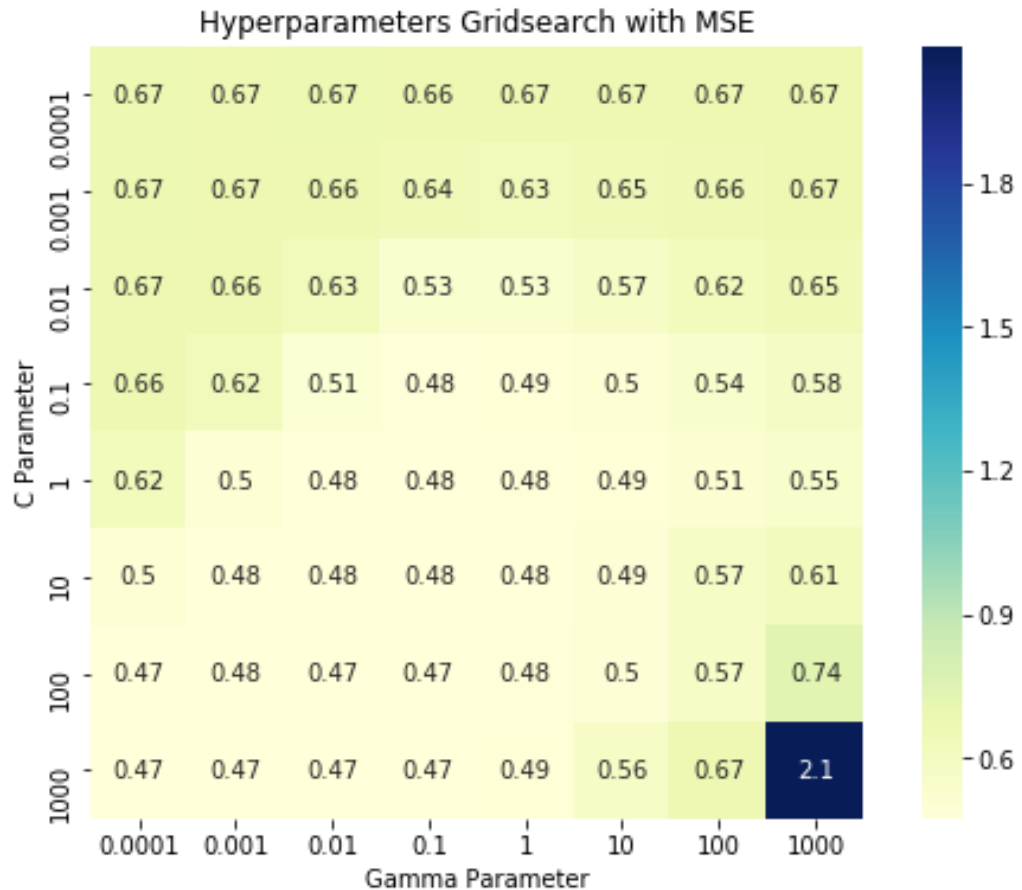
Ans: Accuracy score: 0.8958333333333334

5. (15%) Consider the dataset used in HW1 for regression. Please redo the above questions 2 ~ 4 with the dataset replaced by that used in HW1, while the task is

changed from classification to regression. You should use the [SVM regression model RBF kernel](#) with grid search for hyperparameters and K-fold cross-validation (you can use any K for cross-validation). Then compare the linear regression model you have implemented in HW1 with SVM by showing the Mean Square Errors of both models on the test set.

Ans:

The heatmap of SVR:



I got better MSE from Linear regression than SVM regression model, it seems like Linear regression is more suitable for HW1 data.

Square error of Linear regression: 0.4908933486290753

Square error of SVM regression model: 0.49455182180490137

## Part. 2, Questions (20%):

**The answer of Part 2 is handed written and showed in following pictures.**

- Given a valid kernel  $\kappa_I(\phi, \phi')$ , prove that 1)  $\kappa(\phi, \phi') = \kappa_I(\phi, \phi')$  and 2)  $\kappa(\phi, \phi') = \kappa(\phi) \kappa_I(\phi, \phi') \kappa(\phi')$  are valid kernels, where  $\kappa > 0$  is a positive constant and  $\kappa(\cdot)$  is any real-valued function.

Part 2:

1.

(1) Prove  $k(x, x') = c k_1(x, x')$ , where  $c > 0$

$$\begin{aligned} k(x, x') &= (\nabla_c \phi'_1(x), \dots, \nabla_c \phi'_N(x)) (\nabla_c \phi'_1(x'), \dots, \nabla_c \phi'_N(x')) \\ &= c k_1(x, x') \end{aligned}$$

(2) Prove  $k(x, x') = f(x) k_1(x, x') f(x')$  is valid kernel, where  $f(\cdot)$  is any real-valued function.

Let  $\psi: \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$  be a symmetric positive definite kernel.

Let  $f: \mathbb{R}^N \rightarrow \mathbb{R}$  be any function.

Let  $x_1, \dots, x_M \in \mathbb{R}^N$  and  $c_1, \dots, c_M \in \mathbb{R}$ .

$$\begin{aligned} \sum_{i=1}^M \sum_{j=1}^M c_i k(x_i, x_j) c_j &= \sum_{i=1}^M \sum_{j=1}^M c_i f(x_i) \psi(x_i, x_j) f(x_j) c_j \\ &= \sum_{i=1}^M \sum_{j=1}^M (c_i f(x_i)) \psi(x_i, x_j) (f(x_j) c_j) \end{aligned}$$

So:  $k(x, x') = f(x) \psi(x, x') f(x')$  and because  $\psi(x, x')$  can be another kernel  $k_1(x, x') \Rightarrow$

$$k(x, x') = f(x) k_1(x, x') f(x')$$