

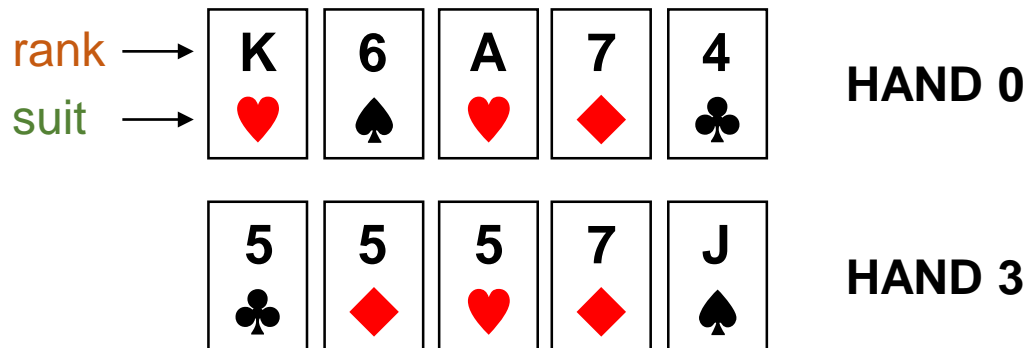
EE412 Introduction to Big Data

Project 1 tutorial

Introduction

- Goal
 - Identify a rule of cards.
- Training dataset
 - consists of 5 cards with (suit, rank).
 - belongs to one of pre-defined classes(called hand).

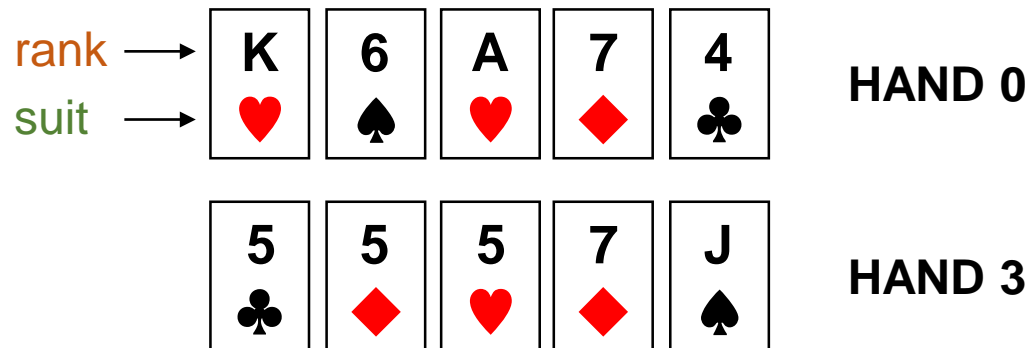
[Examples]



Introduction

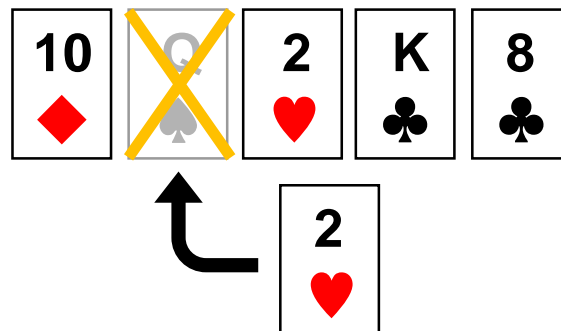
- Suit $\in \{\heartsuit, \spadesuit, \diamondsuit, \clubsuit\}$
- Rank $\in \{2, 3, \dots, 10, J, Q, K, A\}$
- Hand $\in \{0, 1, 2, \dots, 9\}$

[Examples]



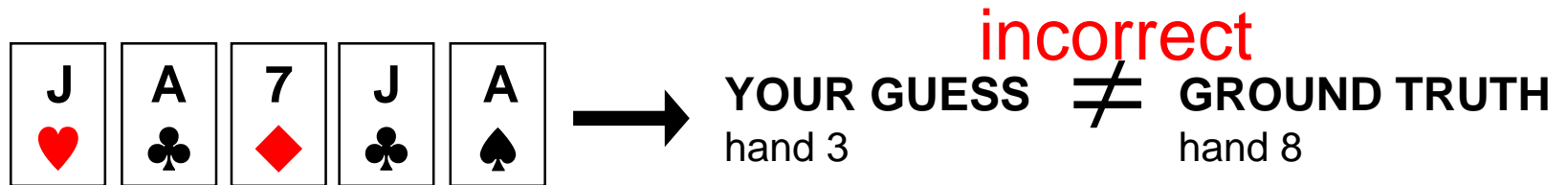
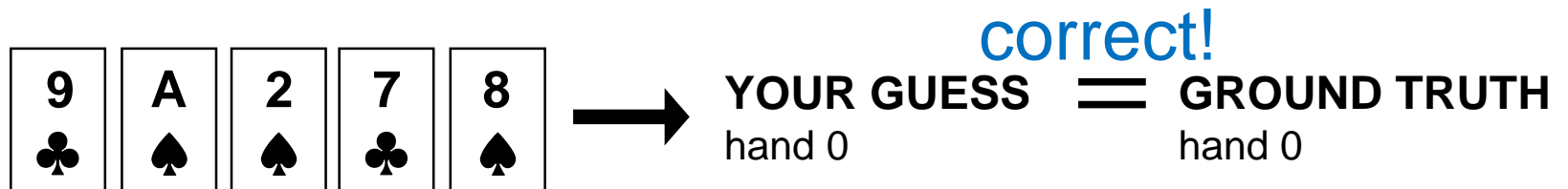
Tasks

- Task 1
 - Classify the hand of cards.
- Task 2
 - Given 5 cards, you can change 1 card into any type of card so that hand of changed cards maximizes.
 - **A new card should not be duplicated.**



Task1: Example

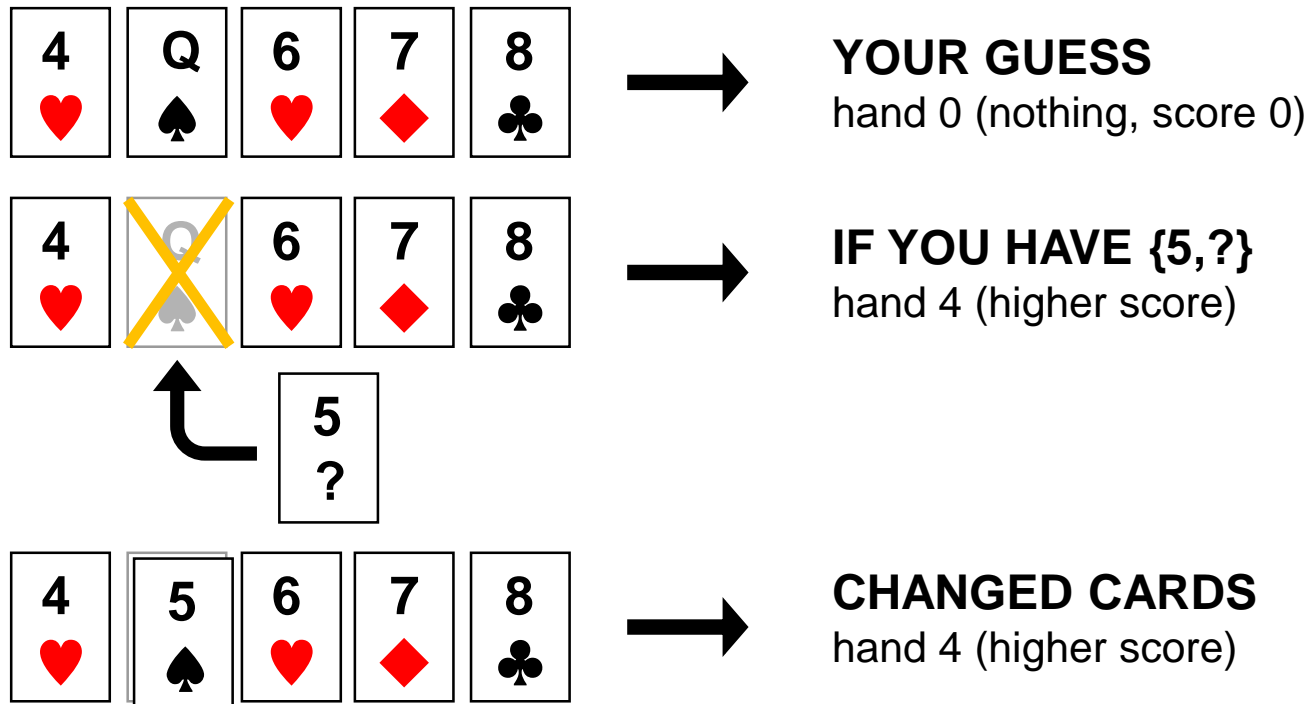
- Goal: classify the hand.



- Evaluated with ratio of correct answer.

Task2: Example

- Goal: make cards with high hand







- Evaluated by score of the modified cards.

Algorithm guide

- Task 1 - classification
 - input: test dataset (csv-format)
 - output: expected hand
- Task 2 – making score high
 - input: test dataset (csv-format)
 - output: modified test dataset

Data format

- Training/test data format
 - csv-format (comma-separated value)
 - each dataset has 11 values: 5 x (suit, rank) + hand
 - Suit of card encodes to {1,2,3,4}
 - 1: , 2: , 3: , 4: 
 - Rank of card encodes to {1,2,...,13}
 - 1: Ace, 11: J, 12: Q, 13: K

[Examples]

S1	R1	S2	R2	S3	R3	S4	R4	S5	R5	hand
1	11	4	1	3	7	4	11	2	1	2

Implementation

- Do your project using MATLAB or Python
- When TA runs your codes
 - should create a file named 'output_taskX'.
 - print each result in a distinguished line.
- No error messages are allowed.

Submission

- Due date: 27th Oct 12:00 pm
- Submit source code with zip file
 - file name: **[student-id].zip**
 - **clarify which files contain either task 1 or 2.**
- Documentation
 - explain what algorithms you use for each task.
 - file name: **doc_[student-id].xx** (pdf, docs, hwp,...)
- EX) 20201818.zip
 - └ task1_20201818.m
 - └ task2_20201818.m
 - └ doc_20201818.pdf

Evaluation

- Accuracy (40%)
 - Task1 (20%), Task2 (20%)
 - **0 if submitted algorithm fails to run.**
 - **0 if result has invalid value.**
- Novelty (50%)
 - 0 if you implement algorithm explicitly with if-else.
- Runtime (5%)
 - 0 if your program runs more than an hour.
- Documentation (5%)

Announcement

- Make a team
 - 1~3 members / team
 - If you have mates, please email to TA (until Sunday)
 - Name and student id of your team members
 - TA: Yunhun Jang (yh.jang@kaist.ac.kr)
 - TA will match a team if you don't have teammates (send me an e-mail if you want me to do for you)