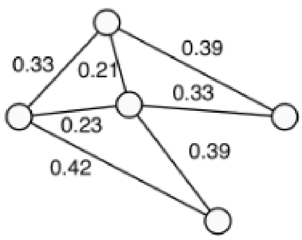
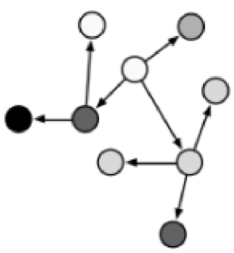
Task 1: Graph vs. Vectorial Pattern Representation

Advantages	Disadvantages
Interpretability: Relationships between entities are explicit and easily understandable, allowing for human inspection and reasoning.	**Scalability:** Processing large graphs can be *computationally expensive*, especially for tasks requiring frequent graph traversal.
Flexibility: Graphs are not constrained to a fixed size and can naturally represent complex relationships and handle different types of entities (e.g., nodes can represent objects, events, or concepts).	**Limited Comparison:** Comparing arbitrary patterns in graphs can be challenging due to the lack of a unified numerical representation.

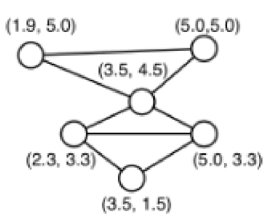
Task 2



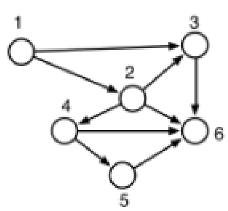
This is a "weighted graph"



This is an "directed graph with labeled nodes" (different shades of grey refer to different labels)



This is a "multiset labeled node graph"

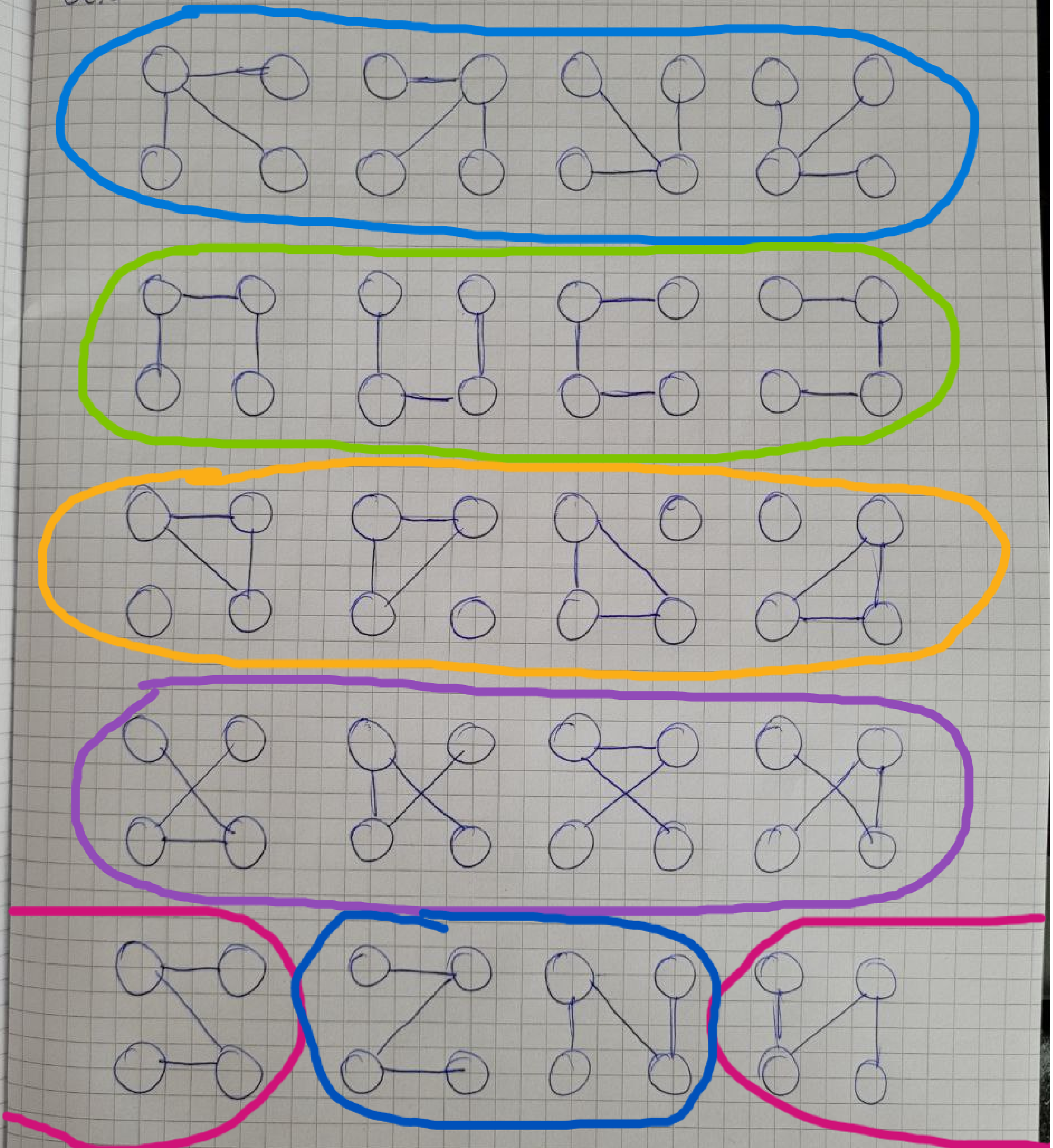


This is a "ordered graph"

Task 3

GBPR

Ex.3



The number of non-isomorphic graphs with n = 4 and k = 3 is N(4,3)=6

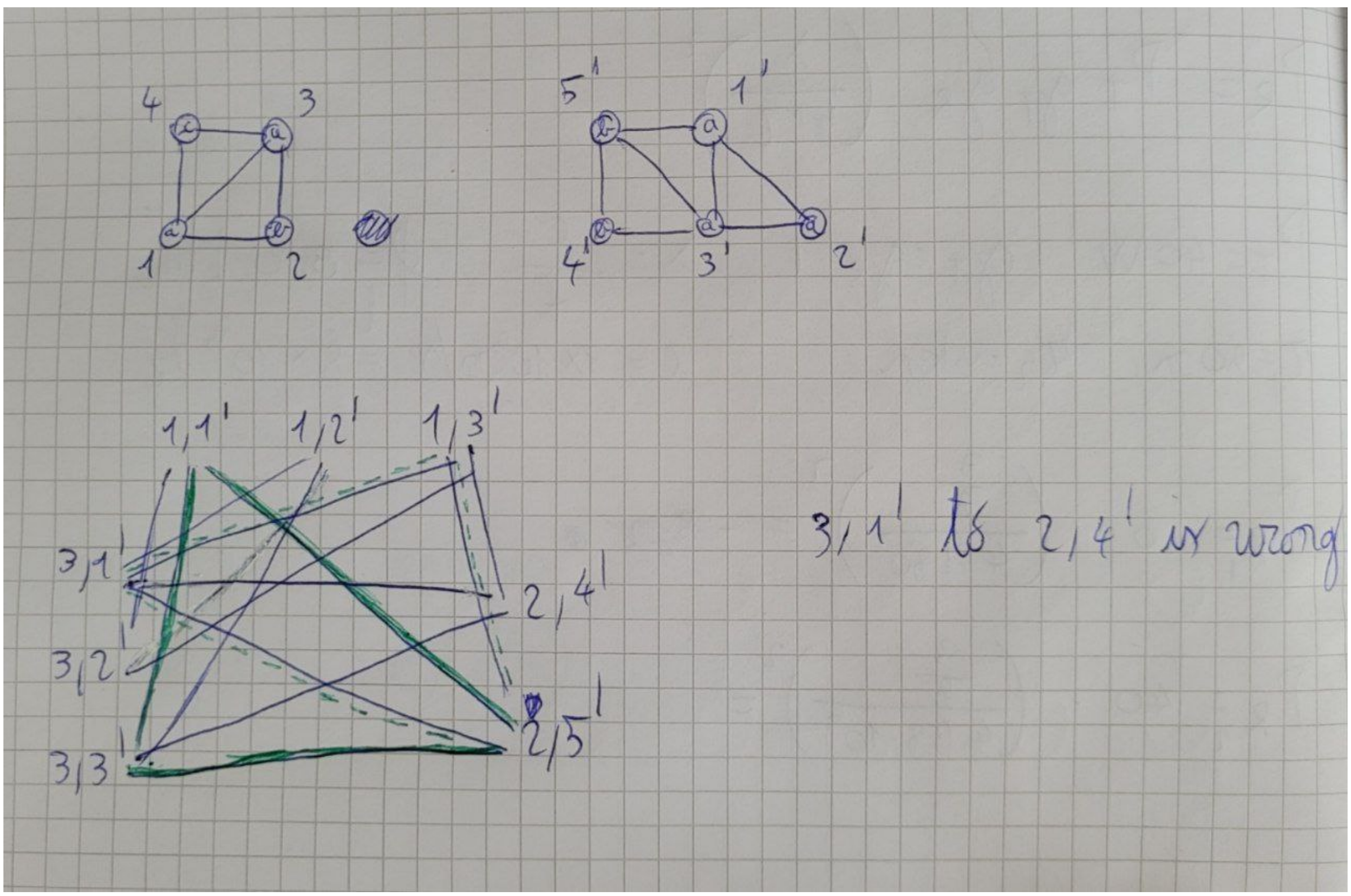
Task 4

The 3 x 3 matrix is represented by each graph with 1 in the entry if two graphs are isomorphic 0 otherwise

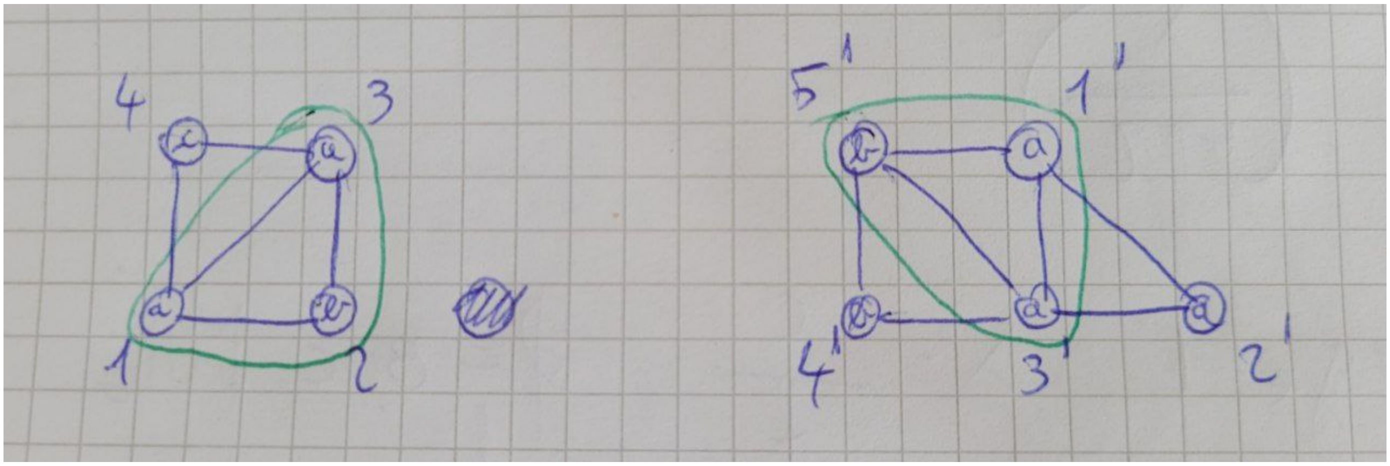
	g1	g2	g3
g1	0	0	1
g2	0	0	0
g3	1	0	0

Explicit isomorphism for graphs g1, g3
f(v1) = x1
f(v2) = x4
f(v3) = x3
f(v4) = x5
f(v5) = x6
f(v6) = x2

Task 5



The maximum clique existing in the association graph are represented by the green lines / dashed line.



This is the maximum common subgraph where we can derive the maximum cliques.

Task 6

a) Shortest-path kernels on graph

Purpose: finding a better algorithm where its time computation is not NP, using graph kernels based on shortest paths.

A graph kernel which is positive definite. Attributes will consist of pairs of the form (attribute-name, value).

Transform the original graphs into shortest-paths graphs. Unlike in the input graph, there exists an edge between all nodes in S which are connected by a walk in I. Every edge between nodes is labeled by
the shortest distance between these two node.

b) Graph Similarity Features for HMM-Based Handwriting Recognition in Historical Documents

Purpose: graph dissimilarity by using the graph edit distance.

The graphs are created without edges where nodes are added to the word graph for each skeleton keypoint and are labelled with its position (x, y). After all keypoints have been included in the word graph, connection points are added.

c) Malware Classification based on Call Graph Clustering

Purpose: use of call graphs to compare the graph similarity scores via graph matching and graph edit distance

A call graph is a directed graph G with vertex set V=V(G), representing the functions (labeled nodes), and edge set E=E(G), where E(G) ⊆ V(G)×V(G), in correspondence with the function calls (directed labeled edges).

d) Graph Matching Based Approach to Fingerprint Classification Using Directional Variance

Purpose: improve the graph edit distance using an efficient algorithm for edit distance computation of planar graphs.

Graph G = (V, E, α, β), where V is the set of nodes, E the set of directed edges, α the node labeling function, and β the edge labeling function.
The neighborhood of a node u in a graph is defined as the induced subgraph N (u) = (Vu , Eu, αu, βu) consisting of node u, all nodes connected to u, and all edges between these nodes.