

# Markscheme

**May 2023**

**Computer science**

**Higher level**

**Paper 3**

© International Baccalaureate Organization 2023

All rights reserved. No part of this product may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without the prior written permission from the IB. Additionally, the license tied with this product prohibits use of any selected files or extracts from this product. Use by third parties, including but not limited to publishers, private teachers, tutoring or study services, preparatory schools, vendors operating curriculum mapping services or teacher resource digital platforms and app developers, whether fee-covered or not, is prohibited and is a criminal offense.

More information on how to request written permission in the form of a license can be obtained from <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

© Organisation du Baccalauréat International 2023

Tous droits réservés. Aucune partie de ce produit ne peut être reproduite sous quelque forme ni par quelque moyen que ce soit, électronique ou mécanique, y compris des systèmes de stockage et de récupération d'informations, sans l'autorisation écrite préalable de l'IB. De plus, la licence associée à ce produit interdit toute utilisation de tout fichier ou extrait sélectionné dans ce produit. L'utilisation par des tiers, y compris, sans toutefois s'y limiter, des éditeurs, des professeurs particuliers, des services de tutorat ou d'aide aux études, des établissements de préparation à l'enseignement supérieur, des fournisseurs de services de planification des programmes d'études, des gestionnaires de plateformes pédagogiques en ligne, et des développeurs d'applications, moyennant paiement ou non, est interdite et constitue une infraction pénale.

Pour plus d'informations sur la procédure à suivre pour obtenir une autorisation écrite sous la forme d'une licence, rendez-vous à l'adresse <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

© Organización del Bachillerato Internacional, 2023

Todos los derechos reservados. No se podrá reproducir ninguna parte de este producto de ninguna forma ni por ningún medio electrónico o mecánico, incluidos los sistemas de almacenamiento y recuperación de información, sin la previa autorización por escrito del IB. Además, la licencia vinculada a este producto prohíbe el uso de todo archivo o fragmento seleccionado de este producto. El uso por parte de terceros —lo que incluye, a título enunciativo, editoriales, profesores particulares, servicios de apoyo académico o ayuda para el estudio, colegios preparatorios, desarrolladores de aplicaciones y entidades que presten servicios de planificación curricular u ofrezcan recursos para docentes mediante plataformas digitales—, ya sea incluido en tasas o no, está prohibido y constituye un delito.

En este enlace encontrará más información sobre cómo solicitar una autorización por escrito en forma de licencia: <https://ibo.org/become-an-ib-school/ib-publishing/licensing/applying-for-a-license/>.

**Subject details:**                    **Computer science HL paper 3 markscheme**

### **Mark allocation**

Candidates are required to answer **all** questions. Total 30 marks.

### **General**

A markscheme often has more specific points worthy of a mark than the total allows. This is intentional. Do not award more than the maximum marks allowed for that part of a question.

When deciding upon alternative answers by candidates to those given in the markscheme, consider the following points:

- Each statement worth one point has a separate line and the end is signified by means of a semi-colon (;).
- An alternative answer or wording is indicated in the markscheme by a “/”; either wording can be accepted.
- Words in ( ... ) in the markscheme are not necessary to gain the mark.
- If the candidate’s answer has the same meaning or can be clearly interpreted as being the same as that in the markscheme then award the mark.
- Mark positively. Give candidates credit for what they have achieved and for what they have got correct, rather than penalizing them for what they have not achieved or what they have got wrong.
- Remember that many candidates are writing in a second language; be forgiving of minor linguistic slips. In this subject effective communication is more important than grammatical accuracy.
- Occasionally, a part of a question may require a calculation whose answer is required for subsequent parts. If an error is made in the first part then it should be penalized. However, if the incorrect answer is used correctly in subsequent parts then **follow through** marks should be awarded. Indicate this with “**FT**”.
- Question 4 is marked against markbands. The markbands represent a single holistic criterion applied to the piece of work. Each markband level descriptor corresponds to a number of marks. When assessing with markbands, a “best fit” approach is used, with markers making a judgment about which particular mark to award from the possible range for each level descriptor, according to how well the candidate’s work fits that descriptor.

**General guidance**

Issue	Guidance
Answering more than the quantity of responses prescribed in the questions	<ul style="list-style-type: none"><li data-bbox="368 333 1485 398">• In the case of an “identify” question read all answers and mark positively up to the maximum marks. Disregard incorrect answers.</li><li data-bbox="368 405 1485 501">• In the case of a “describe” question, which asks for a certain number of facts eg “describe two kinds”, mark the first two correct answers. This could include two descriptions, one description and one identification, or two identifications.</li><li data-bbox="368 508 1485 633">• In the case of an “explain” question, which asks for a specified number of explanations eg “explain two reasons ...”, mark the first two correct answers. This could include two full explanations, one explanation, one partial explanation <i>etc.</i></li></ul>

1 (a) *Award [2 max]*

Public (accept an example, like AWS, Microsoft Azure, Google Cloud);  
Private;  
Community;  
Hybrid;  
Multi-cloud;  
Edge cloud;  
Distributed cloud;

(b) *Award [2 max]*

SaaS provides a software application whereas PaaS provides the environment for software creation/deployment;  
SaaS is meant for consumers whereas PaaS is for developers;  
SaaS is accessed via a web browser/app whereas PaaS can also be accessed through a CLI, SDK, APIs, etc.;  
Data is managed/maintained by the SaaS cloud provider where in PaaS the data is managed/maintained by the client;

2 (a) *Award [4 max]*

Anonymize or pseudonymize data by removing personally identifiable information (e.g. use a lookup table);  
Robust data security measures (e.g., password protect data, encrypt data);  
Transparency/inform users what implicit behavioural data is being collected;  
Obtain users' consent (regarding personalized adverts/sharing data with third parties)/Allow users with options to control their data (i.e. opt out);  
Gather data ethically/do not gather data unethically (e.g., keylogging software/audio recording/video recording);  
Only gather/store data for the required purpose;  
Dispose of data in a timely and secure fashion after stated use has been made;  
Do not share the data with third parties;

(b) **Award [4 max]**

Increase the amount of data/augmentation;  
Track the difference between actual and predicted;  
And stop training before it actual - predicted approaches 0;  
e.g., in k-NN increase k;

Use a dataset with greater variety of data/use multiple datasets/introduce new data over time;  
The model learns to handle a broader range of inputs/learns new features;  
Each dataset has its own biases or noise offering balance;  
And help the model learn the imperfections;

Reduce the complexity of the model;  
By reducing the number of features in the data/choosing only the most crucial features/removing any feature that highly correlates with another;  
Slowly add the features back in while checking for overfitting;  
Use a correlation matrix to calculate the correlation between the features/assess accuracy iteratively;

Early stopping (*Do NOT award reduce training time*);  
Track the model's performance on a validation set;  
Stop when model's performance on validation set degrades (even though training set improves);

Apply regularization parameters;  
Adds a penalty to the loss function;  
To constrain the size of the coefficients/constrain complexity of the model;  
Accept examples of regularisation (e.g. LASSO, RIDGE, Dropout);

Apply cross-validation;  
Divides the dataset into multiple folds;  
Each fold is used as a validation set while the rest are used for training;  
Process is repeated, and the results averaged to estimate the model performance;  
An overfitting model will perform well on validated data but poorly on new data;

Apply ensemble learning;  
Combine multiple models to make a prediction;  
Ensemble has various approaches (e.g. bagging, boosting, and stacking);

Implement Pruning (For decision trees/random forest);  
Pruning involves removing branches or nodes that do not contribute significantly to the model's predictive power;  
Thus, reducing the complexity of the model and the risk of overfitting;

Increase the time gaps (for time-series datasets);  
If data is obtained every minute sampling every 15 minutes leads to fewer data points;  
This avoids adjacent data points that are highly correlated;  
But this only works when you don't have sparsity of data;

**Mark [2] + [2]**

**3 Award [6 max]**

*Award up to [1 max] for a definition*

F-measure is a useful metric to determine the accuracy of a recommender system;  
 Precision is the ability to make correct relevant predictions/recommend suitable content;  
 Recall is what percentage of the relevant predictions from all predictions/how many selected people the system recommended;

*Award up to [5 max] explaining the process*

Loop through the 20 user clips;  
 Check if the user's clip was recommended or not/Check prediction;  
 Check if the user was given a role or not/Check actual;  
 Tag the clip for accuracy (i.e. True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN));  
 Check the system's ability to make correct recommendations/recommend suitable content/determine precision;  
 Check the system's ability to recommend users' clips that resulted in a role/determine recall;  
 Calculate the harmonic mean of precision and recall/F-measure formula (see below);  
 Allow the user to add a weighting to either precision or recall/Apply a value for  $\beta$ ;

$$F_{\beta} = (1 + \beta^2) \times \frac{\text{precision} \times \text{recall}}{(\beta^2 \times \text{precision}) + \text{recall}}$$

$$= \frac{(1 + \beta^2)tp}{(1 + \beta^2)tp + \beta^2fn + fp}$$

Precision = TP / (TP + FP) or 3 / (3 + 2) = 60% or 3/5;

Recall = TP / (TP+FN) or 3 / (3 + 1) = 75% or 3/4;

F-Measure (F1) = 67% or 2/3 (calculation not needed – state combine precision and recall);

*Award up to [2 max] for explaining the purpose*

By calculating the F-score the system can be compared with other systems/settings;  
 Every time the hyperparameters are changed the F-scores can be compared;  
 To identify which system/settings is better suited for the application;

#### 4 Award [12 max]

Answers may include:

##### **Content-based filtering**

- A content-based approach requires a good amount of information about items' features (rather than users' interactions).
- They can genre, year, performer, textual information (extracted using NLP).
- Uses item's features to recommend items similar to what the user likes.
- Content-based filtering does not use other users' data for making recommendations.
- The filtering system builds up a user profile.
- The user's likes are determined by behaviours;
- They can be explicit (e.g, adding a rating) or implicit (clicking on a link to another video by the same artist).
- The content is rated by many users and has a mean score.
- New users or items can be described by their characteristics (e.g., in your profile you may indicate a preference for rap so any content in this genre would score highly and be recommended).
- Content-based filtering is dependent on Item Representation: The effectiveness of these systems largely depends on how items are represented, and they may not work well if the item descriptions are not accurate or comprehensive.
- Content-based filtering provides limited diversity for recommendations.
- Content-based filtering copes with a cold-start but will need some initial starting data (e.g. features and/or ratings input into the system).

##### **Collaborative filtering**

- Collaborative filtering is a method of making automatic predictions about the interests of a user by collecting preferences from many users.
- Collaborative filtering only needs a user's historical preference on a set of items to work.
- The core assumption is that users who agreed with each other in the past will agree in the future.
- Content knowledge (e.g., features) is not needed/does require explicit labelling of data.
- CF requires large datasets to be effective.
- CF systems can also be content-aware.
- Context-aware collaborative filtering systems have to deal with a larger dataset.
- Matrix factorization is an effective model for handling the larger dataset and making recommendations.
- CF has a cold start problem and may struggle to make recommendations for new users and new items due to a lack of data/lack of user interactions/no preferences for new users.
- CF suffers from data sparsity problem so different techniques are needed.
- CF is based on the premise that users (preferences) and items have been profiled correctly but behaviour and ratings may change in the future.
- CF suffers from poor scalability because computations are expensive and grow non-linearly when number of users and items added.
- CF suffers with popularity bias because popular items have more interactions, are more likely to be recommended, which leads to a feedback loop.
- CF may struggle with heterogeneous (diverse) populations or unique individuals/works better with homogenous populations because recommendations assume people have similar interests.
- CF can recommend items that are quite different from those the user has previously consumed, potentially introducing users to new interests.

### Comparison between content-based and collaborative filtering systems

- Content-based does not suffer from the cold start of collaborative filtering systems (need many users' preferences before they become effective).
- Knowledge-based techniques can address the cold start problem/Build up a profile of the users (e.g. demographics).
- Content-based does need the features to be described (e.g., genre, artist) whereas collaborative filtering does not need any features.
- Content-based filtering does not data about other users because the recommendations are specific to one user whereas collaborative filtering requires all user's preferences.
- Collaborative filtering helps users discover new interests by association (another user explored new content and liked it, their interests are similar to yours, so you may like this) whereas Content-filtering will only recommend more of what the user already likes.
- Content-based filtering must store data on users (even though different users' data is not shared) and items whereas collaborative filtering only needs user data (i.e., preferences/behaviour).
- Unlike Content-based, Collaborative filtering has not need for Item Representation. It does not require a detailed description or representation of each item.

### Hybrid content-based / collaborative systems

- This approach would offer the advantages of both systems (see comparison).
- Would be an effective approach especially when there is limited user data (early in the business venture).
- Increased complexity and more difficult to implement.
- Might be sensible to start with a content-based filter then transition to a collaborative or hybrid model as *NextStar* becomes more popular.

### Supervised learning

- Labelled datasets are used to train the model.
- The system learns to map an input to an output based on examples.
- A trained supervised learning system produces an inferred function which it uses to classify new data.
- Thus, the model can generalize within confined parameters.
- The statistical quality of the model can be measured.
- There are numerous types of supervised learning algorithms. Each is suited to particular circumstances like the quality and type of data.

### Unsupervised learning

- Learns patterns from untagged (unlabelled) data.
- Unsupervised models exhibit self-organization that captures patterns as probability densities.
- Two broad approaches are neural networks and probabilistic methods.
- an unsupervised neural network tries to mimic the data it is given and uses the error (difference) in its mimicked output to correct itself by adjusting its weights & biases).

### Reinforcement learning

- A method based on rewarding desired behaviours/punishing undesired behaviours.
- Takes action to maximize cumulative reward (exploitation) and balances against random actions (exploration).
- Does not require explicit correction of sub-optimal actions.
- Learns through trial and error.
- Modelled as a Markov decision process.

**Hybrid model**

- Extremely complex to implement.
- It can take the best aspects of each learning approach so if implemented effectively it will produce the best model.
- Supervised learning takes advantage of training so that the model produces good recommendations early.
- Unsupervised learning may find unexpected features that a human might miss, thereby producing better recommendations.
- The feedback aspect of reinforcement learning allows the system to be informed of bad decisions, adjust and improve.

**Conclusion.**

- A final measured conclusion is included in which the candidate links together the various points.

*Please see markband on page 11.*

Marks	Level descriptor
No marks	<ul style="list-style-type: none"> <li>No knowledge or understanding of the relevant issues and concepts.</li> <li>No use of appropriate terminology.</li> </ul>
Basic 1-3 marks	<ul style="list-style-type: none"> <li>Minimal knowledge and understanding of the relevant issues or concepts.</li> <li>Minimal use of appropriate terminology.</li> <li>The answer may be little more than a list.</li> <li>No reference is made to the information in the case study or independent research.</li> </ul>
Adequate 4-6 marks	<ul style="list-style-type: none"> <li>A descriptive response with limited knowledge and/or understanding of the relevant issues or concepts.</li> <li>A limited use of appropriate terminology.</li> <li>There is limited evidence of analysis.</li> <li>There is evidence that limited research has been undertaken.</li> </ul>
Competent 7-9 marks	<ul style="list-style-type: none"> <li>A response with knowledge and understanding of the related issues and/or concepts.</li> <li>A response that uses terminology appropriately in places.</li> <li>There is some evidence of analysis.</li> <li>There is evidence that research has been undertaken.</li> </ul>
Proficient 10-12 marks	<ul style="list-style-type: none"> <li>A response with a detailed knowledge and clear understanding of the computer science.</li> <li>A response that uses terminology appropriately throughout.</li> <li>There is competent and balanced analysis.</li> <li>Conclusions are drawn that are linked to the analysis.</li> <li>There is clear evidence that extensive research has been undertaken.</li> </ul>

**Total: [30]**

---