# What Song Am I Thinking Of?

Niall McGuire and Yashar Moshfeghi

NeuraSearch Laboratory, Department of Computer and Information Sciences, of
Strathclyde, Glasgow, Scotland niall.mcguire@strath.ac.uk
yashar.moshfeghi@strath.ac.uk

**Abstract.** Information Need (IN) is a complex phenomenon due to the difficulty experienced when realising and formulating it into a query format. This leads to a semantic gap between the IN and its representation (e.g., the query). Studies have investigated techniques to bridge this gap by using neurophysiological features. Music Information Retrieval (MIR) is a sub-field of IR that could greatly benefit from bridging the gap between IN and query, as songs present an acute challenge for IR systems. A searcher may be able to recall/imagine a piece of music they wish to search for but still need to remember key pieces of information (title, artist, lyrics) used to formulate a query that an IR system can process. Although, if a MIR system could understand the imagined song, it may allow the searcher to satisfy their IN better. As such, in this study, we aim to investigate the possibility of detecting pieces from Electroencephalogram (EEG) signals captured while participants *"listen"* to or *"imagine"* songs. We employ six machine learning models on the publicly available data set, OpenMIIR. In the model training phase, we devised several experiment scenarios to explore the capabilities of the models to determine the potential effectiveness of Perceived and Imagined EEG song data in a MIR system. Our results show that, firstly, we can detect perceived songs using the recorded brain signals, with an accuracy of 62.0% (SD 5.4%). Furthermore, we classified imagined songs with an accuracy of 60.8% (SD 13.2%). Insightful results were also gained from several experiment scenarios presented within this paper. Overall, the encouraging results produced by this study are a crucial step towards information retrieval systems capable of interpreting INs from the brain, which can help alleviate the semantic gap's negative impact on information retrieval.

**Keywords:** Information systems · Information retrieval · Music Retrieval · Brain · EEG · Machine Learning.

## 1 Introduction

All Information Retrieval (IR) systems aim to satisfy searchers' Information Needs (IN). Many IR systems rely on the searcher to assess their information requirements and convey them to the system, typically in the form of queries [58]. However, conventional queries may suffer from semantic gaps between the searcher's true IN and their formulated query [54, 17]. Various studies [34] have tried to alleviate this problem by designing an IR system capable of understanding INs from biological features. Studies such as [36] focus on the use of Functional Magnetic Resonance Imaging (fMRI) to analyse the Blood Oxygenation Level Dependent (BOLD) signals to detect which areas of the brain are more active in the process of realising IN. Alternative methods to fMRI, such as Electroencephalograms (EEG), have been employed to understand further how INs are realised within a subject [28].

A subdivision of the IR field that may greatly benefit from the development of an IR system that can interpret INs from the brain is Music Information Retrieval (MIR) [46]. MIR is exploring and organising enormous amounts of music or music information based on its relevance to specific queries [46]. Recently this area has seen the use of EEG data to extract music-related information from subjects [30, 15, 49]. Prior studies have shown the ability to extract the tempo and meter of the music stimuli subjects are exposed to from their EEG data [50]. At the same time, others have successfully utilised EEG data that was recorded while a subject listened to a segment of music. To then determine which song the subject perceives solely from their brain signals.

Although these prior studies have demonstrated the potential of **Perceived** EEG song data for extracting music-related information [51, 49], they have yet to effectively demonstrate the ability of **Imagined** EEG song data to extract song-related information for potential use within an IR system, which could then be consolidated into a MIR system, allowing for more efficient satisfaction of a searcher's IN. Thus, to start building towards a brain-driven MIR system, in this study, we aim to answer the following research questions:

- **RQ1**: "*Is it possible to classify a song from subjects' EEG data whilst Perceiving a given song?*"
- **RQ2**: "*Is it possible to classify a song from subjects' EEG data whilst Imagining a given song?*"
- **RQ3**: "*How similar are Perceived and Imagined EEG data? and can they be used interchangeably for training and testing of machine learning models*"
- **RQ4**: "*How much of a song would a subject have to Perceive/Imagine before the system can produce accurate results?*"

## 2 Background

### 2.1 Information Need Complexity

Information need (IN) is one of the key concepts in IR theory [24]. The subjective nature of IN makes it complex for IR systems to fully comprehend [44]. The source of this complexity lies in the IN's paradoxical nature, meaning that unlike many other essential human needs such as hunger or fatigue, the information that a subject requires to satisfy their needs is most likely unknown to the searcher. In prior studies, this has been explained by the idea that an IN is "intangible and visceral" and thus is "unknowable and non-specifiable", this creates a conundrum for the searcher such that if they are unable to create a query proficient enough for an IR system to understand and retrieve relevant documents, then they are unable to satisfy their IN [4, 7]. An IN can link to a discrepancy in knowledge [2, 3], uncertainty [61, 26], anxiety [8], dissatisfaction [52], or doubt [54]. Many researchers have attempted to better understand this complex phenomenon in the past [2, 26, 53]. Searchers will realise an internal IN when presented with a gap within their Anomalous States of Knowledge (ASK) [4]; this, in turn, sets the search process into motion. A searcher would begin this process by transposing their IN into a suitable query processable by an IR system, this then prompts the IR systems to retrieve documents likely relevant to the initial query to fulfil the searchers IN [8], followed by the searcher reviewing the documents presented by the IR system and taking in any relevant information to satisfy their IN [36]. Unfortunately, it is common that when reviewing the output of the IR system, searchers will find that their IN is not sufficiently satisfied [43] firstly, since the IN was not formulated correctly. Secondly, due to the transformation of the IN into a query represented by keywords, which are considered to be noisy and uncertain [17] as they are limited as to how close they can be to the IN [54].

### 2.2 NeuraSearch

Recently, studies have begun utilising neurophysiological features from searchers in an attempt to understand IN information, this area of research has gone under the term NeuraSearch [28, 31]. The widely utilised neuro-imaging modalities for this purpose include EEG [41, 37, 1, 13, 18, 20, 32, 34, 33, 57, 35], Functional Magnetic Resonance Imaging (fMRI) [32, 34, 37, 39, 40], Magnetoencephalography (MEG) [23], Functional Near Infrared Spectroscopy (fNIRS) [27, 29]. Among these powerful techniques, EEG has drawn more attention from researchers because of its high temporal resolution (millisecond scale) and unobtrusiveness. EEG can be deployed in both clinical settings, such as to predict patients' epileptic seizure [56], stroke [19], or brain tumour [22], and non-clinical settings, including detecting fatigue [11], or performing a recognition task [9] or a realisation of IN [10]. Research conducted by [28, 6, 21] saw the utilisation of EEG, which has a high temporal resolution that allows for real-time capture of brain signals and Event-Related Potentials (ERP) to observe the brain whilst participants developed an IN. It can facilitate information acquisition performance [60]. The findings of this study [59] showed that the realisation of an IN is developed within the brain before the consciousness of the subject observes it. With the development of the understanding and processing of INs from neurophysiological features [48, 55], the possibility of using said features within practical applications to overcome the issue of noisy and ill-defined IN queries becomes even more feasible [1, 28]. A core area within IR that suffers substantially from the use of keyword queries is that of MIR. This is because current MIR systems require users to present their query using metadata such as the song title, lyrics, or the artist's name [46]. However, if the IN needs to be defined and the searcher cannot recall these pieces of information, they may have to imitate the desired song through singing/humming [51]. As a result, MIR may benefit significantly from an IR system capable of processing internally imagined song information as a query and obtaining relevant documents to the imagined song query [42].

### 2.3 Neuroscience and MIR

Recent works have begun incorporating neurophysiological data with Music Information Retrieval (MIR) applications. An early example of using EEG within MIR can be found in [45], where researchers had 10 participants listen to 7 short melody segments ranging from 3.26 to 4.36 seconds. A logistic regression classifier was utilised to classify the ERP of each trial. This study showed an above random classified accuracy that, when applied to individual subjects, varied from 25% to 70%, whereas when applied across all subjects, it ranged from 35% to 53%. Within works by [51], the researchers created the OpenMIIR data set, which contains the EEG data of participants who were recorded while exposed to several short music segments, as well as the recordings of those same participants asked to then imagine the song segment. This data set was then used in a follow-up study by [50], in which the researchers attempted to classify which song the participants were perceiving/imagining. A Support Vector Machine (SVM) and a Neural Network (NN) were used to classify the music from the perceived EEG data. The results showed that the SVM had a classification accuracy of 27.59% for the perceived data, and the NN had a classification accuracy of 27.22%, which is significantly higher than random classification (8%). In a more recent study by [38], the authors use the perceived EEG song data from the OpenMIIR data set to train various machine learning models on 9 out of the ten subjects, where the last subject was used for evaluating the models. The results from their study presented a song classification rate varying from 52.5% to 24%. Although these previous studies have achieved accuracy scores substantially higher than random classification (8%), there is still room for significant improvement in overall classification accuracy and prediction variance. As well as this, there has been no reporting of the above random classification of the Imagined song segments. This is arguably the most significant part of creating a MIR system capable of interpreting IN directly from a subject's brain hence why the findings presented within this paper are of great significance to tackling this problem.

## 3 Methodology

The main steps of this study are EEG signal acquisition, artefact removal, feature engineering, model evaluation and EEG music classification. Details of each step will be described in the following sections.

### 3.1 Data set

Within this study, the OpenMIIR dataset was used to provide subject EEG data [51]. The dataset contains the EEG signals of 10 subjects recorded using 64 + 2

EEG electrodes sampled at a frequency of 512 Hz. As well as this, the horizontal and vertical EOG channels were recorded to capture the eye movements of the subject. The dataset was produced by having each subject perceive and imagine 12 short pieces of music 5 times to create repetitions. Each song ranged between 7 to 16 seconds in length whilst their EEG data was recorded. The short music segments are taken from a variety of musical pieces that varied across musical genres, songs that were recorded without lyrics, songs where the lyrics have been removed and purely instrumental pieces. Although the original dataset was captured under four various conditions, we only make use of one of those four conditions for this study, this being the recording of the subject's EEG data whilst they are perceiving a short music segment.

### 3.2 EEG Preprocessing

During EEG recording it is commonplace for electrical activities caused by the actions of the individual to affect the data causing artefacts to form. These are unwanted signal fragments that can negatively impact measurements and skew results. Thus, it is a crucial step to remove these artefacts to the best of our ability, to achieve this we made use of the MNE python toolbox [12]. Firstly, "bad" channels that had previously been marked by the dataset creator were removed from the raw dataset. This was then followed by performing channel interpolation.

Often during EEG recording the skin–electrolyte–electrode interface produces a drift in voltage which can be picked up by the EEG electrodes [16], to account for this, we filtered out these drifts using EEG using a bandpass filter with a frequency ranging from 0.5 Hz to 40 Hz. Another artefact that may be present in the data is the electrical interference caused by the subject blinking their eye, to handle this, we applied the Independent Component Analysis (ICA) technique which computes the independent components within the raw dataset allowing us to remove any components that share a high correlation with the EOG channels. The 64 EEG channels were then reconstructed from the remaining independent components.

### 3.3 EEG Feature Extraction

To capture significant EEG signal characteristics, we generated features from the processed EEG data. The details of features in their categories are discussed below:

1. **Morphological features:** the number of peaks, average non-linear energy and the curve length were extracted. Features and three morphological features were extracted in two different ways. Firstly, they were calculated at four frequency bands theta (4–8 Hz), alpha (8–13 Hz), beta (13–25 Hz), and low gamma (25–40 Hz).
2. **Statistical features:** skewness, mean, and kurtosis were computed to identify the distribution of the signal.[25] was calculated to describe time-varying processes

3. **non-linear features:** The approximate entropy (ApEn) and Hurst exponent (H) were used to quantify the unpredictability of fluctuations over a time series as well as measure the self-similarity of the time series, respectively.

### 3.4 Models

Machine learning techniques, particularly advanced models such as deep learning, have been employed to accurately extract variance characteristics in EEG data for song identification[47]. Therefore, to investigate the research questions, we adopted the six machine learning models: GNB, SVM, GRU, BGRU, LSTM, and BLSTM, in our analysis. These models have been widely used in the EEG signal for classification [5, 25]. GRU, BGRU, LSTM, and BLSTM are beneficial for learning sequential data with long-term dependencies [62].

**Table 1.** Deep learning model architectures

| Model | Layers/Nodes |
|---|---|
| GRU | G256-G128-D12 |
| BGRU | BG256-G128-D12 |
| LSTM | L256-L128-D12 |
| BLSTM | BL256-L128-D12 |

The architectures of the deep learning models are shown in Table 1. G, BG, L, BL, and D correspond to GRU, BGRU, LSTM, BLSTM, and Dense layer, respectively. For example, BL256-L128-D12 implies that there is a BLSTM layer with 256 units, an LSTM layer with 128 units and since there are twelve classes in our analysis, a Dense layer with twelve outputs for each model. Softmax activation was implemented in the last layer. This work implemented a dropout rate of 0.2, and Adam was used to train all deep-learning models. Additionally, early stopping was utilised to prevent an over-fitting issue. Furthermore, we stopped training after the model's performance stopped increasing after 200 epochs.

### 3.5 Classification Scenarios

To address each research question, we created several experimental scenarios, which are detailed as follows:

**Scenario 1: Perceived/Imagined song classification** To investigate **RQ1** and **RQ2**, our model performs classification utilising the EEG signals of all subjects, bar one subjects data, that was left out of the training set and used for testing purposes, creating a 10-fold leave one out cross validation training loop (LOOCV). The task was performed twice using two distinct groups of data;

perception (**RQ1**) and imagination (**RQ2**), and classification was performed for each repetition in each stage to determine the difference in model performance (**RQ3**).

**Scenario 2: Interchanging of Perceived and Imagined EEG music classification** To investigate **RQ3** and determine whether imagined and perceived songs can be classified interchangeably. Our models were trained with EEG signals from the perception stage and tested with EEG signals from the imagination stage, and vice versa. Like the first experiment condition, we again rotated the subject whose data was used to evaluate the model for both the perceived and imagined states creating a 10-fold LOOCV loop.

**Scenario 3: Train/Test data limitation** To investigate **RQ4**, we trained the models the same way as in Scenario 1, but how we test our models is different. After completing the model training procedure, we test our model utilising n% of the initial portion of the data for each song repetition. The sizes of the testing data set varied between 10%, 25%, 50%, 75%, and 100%.

### 3.6 Metrics

This study evaluates the model performance using accuracy, precision, and recall.

**Accuracy** is the sum of the number of true positives and true negatives divided by the total number of examples.

**Precision** refers to the ratio of correct positive examples to the number of actual positive examples.

**Recall** is calculated as the sum of true positives across all classes divided by the sum of true positives and false negatives across all classes.

## 4 Results

In this section, we use a box plot to display the results for the models in several experimental scenarios. Over the ten cross-validation sets, each box plot presents five crucial pieces of information: the minimum, first, second (median), third, and maximum quarterlies [14], where the red dot in the box represents the mean of accuracy for each model.

### 4.1 Perceived/Imagined song classification

The results from training the models on the Perceived and Imagined datasets are presented in Figure 1 and Figure 2. Figure 1 details that the highest performing model trained on Perceived EEG song data is the LSTM at the fourth repetition with an average accuracy score of 62.0% (SD 5.4%). This result helps to address **RQ1**, demonstrating the ability of the techniques employed within this study to provide a higher classification accuracy than that achieved by prior studies.

As shown in Figure 2, the highest performing model is again the LSTM on the fifth repetition with an average accuracy score of 60.8% (SD 13.2%). The result achieved by the LSTM trained on Imagined EEG song data helps to satisfy **RQ2**. Firstly, by showing that the techniques employed within this study can classify songs solely from Imagined music EEG data. Secondly, the model achieved a significantly higher classification score than random classification (8%). As well as, this Imagined classification was within 2% of Perceived classification, demonstrating the comparable performance of Imagined EEG data with Perceived EEG data. Further observations of both figures show that each model achieves similar results on both Perceived and Imagined data with slight variations. GRU, LSTM, BRGU, and BLSTM's accuracy appears to increase with the incremental addition of each repetition into the training data set, except GNB and SVM.

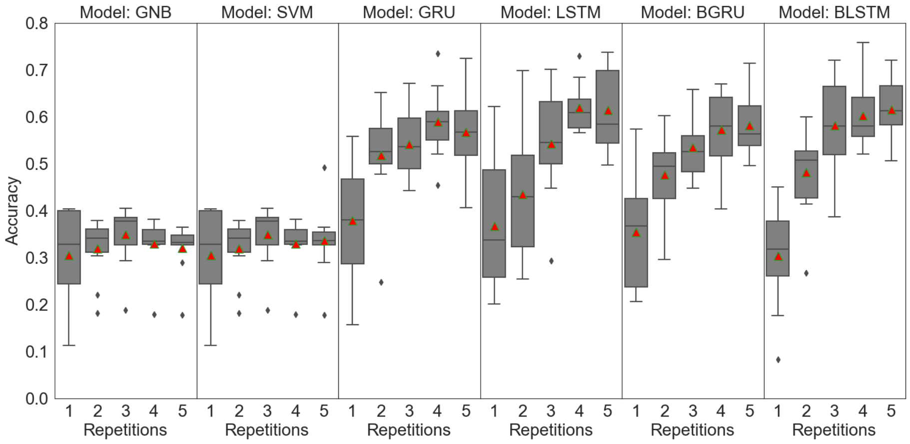

**Fig. 1.** Average results of models trained on **Perceived** song data

### 4.2 Testing Size Breakdown

The results of this experiment are presented in Figure 3 and Figure 4. The results illustrated in both figures highlight that the highest performing models at the smallest test sample size of 2 (which translates to 10% of the comprehensive testing data) are the Perceived GNB/SVM achieving an accuracy of 30.3% (SD 6.6%) and the Imagined GNB/SVM both scoring 30.5% (SD 5.2%). However, their accuracy falls behind the other models after the initial testing segment. Interestingly, when observing the performance of GRU, LSTM, BGRU, and BLSTM across both data types. We can see that segment sizes 10 (50% of testing) and 15 (75% of testing) share similar accuracy scores; for example, the highest Perceived model at segments 10 and 15 is the BLSTM with 57.3% (SD 5.5%) and 56.5% (SD 5.4%) respectively. Similarly, the highest Imagined model

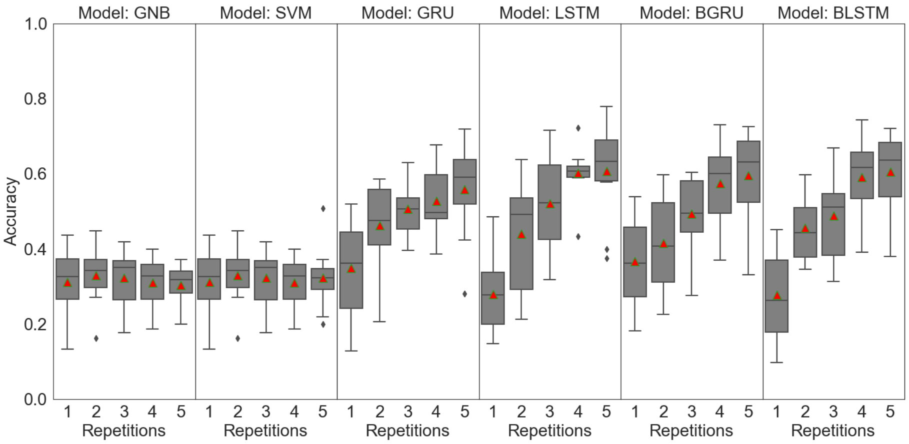

**Fig. 2.** Average results of models trained on **Imagined** song data

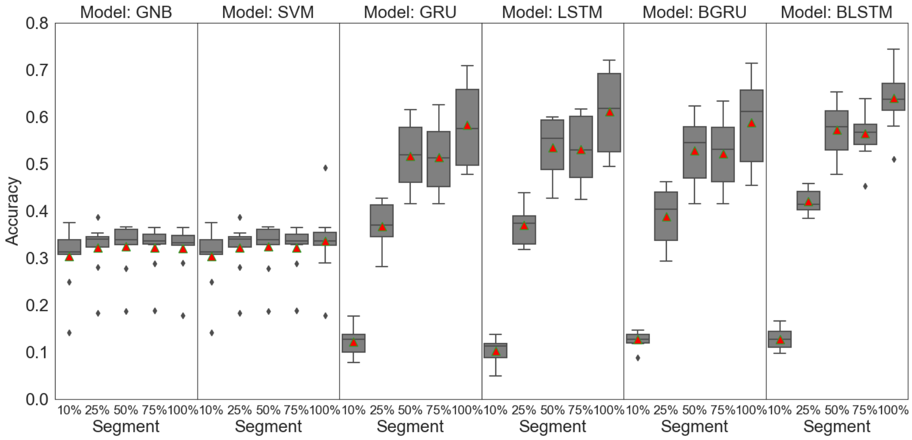

**Fig. 3.** Average results of models tested on the varied **Perceived** data set

at segments 10 and 15 is the LSTM which achieved 52.5% (SD 8.2%) and 53.4% (SD 7.3%). As well as this, the increase from segment size 10 to segment size 20 (100% of testing) does increase the accuracy significantly, with the Perceived BLSTM scoring accuracy of 64.1% (SD 6.9%) and the Imagined LSTM scoring 61.8% (SD 8.7%) on segment 20 (100% of the testing set). An important consideration when viewing these results is that the song segments used for training already average around 7's in length. That means if we use 50% of the test set and achieve a classification close to 50%, we are achieving that by using only 3.5's of an entire song. These results help to address **RQ4**.

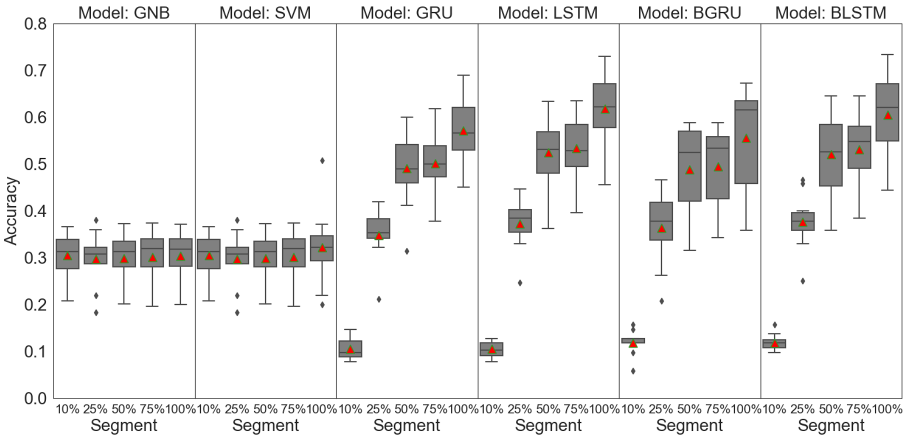

**Fig. 4.** Average results of models tested on the varied **Imagined** data set

**Table 2.** Interchanging of Perceived and Imagined for train/test

| Trained on/Tested on | Model | Mean Precision (SD) | Mean Recall (SD) | Mean Accuracy (SD) |
|---|---|---|---|---|
| Perceived/Imagined | GNB | 31.8%, (5.9%) | 32.0%, (6.8%) | 32.0%, (6.8%) |
| | SVM | 43.7%, (12.7%) | 43.7%, (12.4%) | 43.7%, (12.4%) |
| | GRU | 59.2%, (7.4%) | 60.0%, (9.0%) | 60.0%, (9.0%) |
| | LSTM | 62.1%, (8.5%) | 63.4%, (9.4%) | 63.4%, (9.4%) |
| | BGRU | 59.7%, (10.6%) | 59.9%, (10.9%) | 59.9%, (10.9%) |
| | BLSTM | 62.1%, (8.5%) | 63.6%, (8.8%) | 63.6%, (8.8%) |
| Imagined/Perceived | GNB | 25.6%, (6.2%) | 19.1%, (5.6%) | 19.1%, (5.6%) |
| | SVM | 34.5%, (4.3%) | 34.0%, (4.6%) | 34.0%, (4.6%) |
| | GRU | 57.4%, (8.8%) | 57.5%, (8.9%) | 57.5%, (8.9%) |
| | LSTM | 61.2%, (8.3%) | 61.4%, (7.5%) | 61.4%, (7.5%) |
| | BGRU | 54.6%, (7.9%) | 56.3%, (8.7%) | 56.3%, (8.7%) |
| | BLSTM | 61.2%, (8.3%) | 61.2%, (8.2%) | 61.2%, (8.2%) |

### 4.3 Changing of Perceived and Imagined for train/test

The results of this scenario are detailed in Table 2 for the initial case of training the models on the perceived data set and evaluating them on the imagined data. We can observe that the highest performing model for this case is the BLSTM with a score of 63.6% (SD 8.8%). Furthermore, in the case where we trained on the imagined data and evaluated the perceived data. The model with the highest accuracy is the LSTM, with a score of 63.4% (SD 7.5%). The results highlight the comparable nature of Perceived and Imagined EEG song data. The findings of this experiment scenario help to directly address **RQ3**.

## 5 Discussion and Conclusion

The findings from our first experiment scenario, as seen in Figure 1 and Figure 2, directly address **RQ1** and **RQ2**. These results give significant weight to the points made earlier in the introduction on the effectiveness of Imagined EEG data to extract IN-related data that can then be used to provide the searcher with documents capable of satisfying their IN. In our case, we have successfully extracted a potential IN (an Imagined song) directly from neurological features (recorded EEG data) and created a system capable of interpreting those features into data that a standard IR system could then understand. From these initial findings, we can also see the similar nature of Perceived and Imagined data. Although the EEG data for Perceived and Imagined are collected under two different conditions, the models appear to share performance across both types, this could imply that the process of Perceiving and Imagining a song follows a similar process within the brain and produces a pattern that the machine learning models can identify.

The second experiment scenario was designed to address **RQ3**. The results from this experiment were displayed in Table 2. We observe that the models can classify songs from a data type (e.g. Perceived or Imagined) that it had yet to be trained on, at a similar classification rate across both data types. The models produced classification results comparable to those presented in the first experiment scenario when the models were trained and evaluated on the same data type, this adds further weight to the point that Perceiving and Imagining a song can produce similar EEG outputs that the models can interpret. This implies that for a real-world brain-driven MIR system, that no matter what data type the searcher presents to the system, it will likely be capable of classifying the song correctly.

Experiment scenario three results are displayed in Figures 3 and 4. The findings from this experiment highlighted the ability of the models to detect the songs with 50% accuracy when given only 3.5s of the music segment, this has significant implications for a brain-driven MIR system as it would be the system's goal to limit the time a searcher would have to Perceive/Imagine a song while producing the most optimal classification results.

In conclusion, this study investigated the use of neurophysiological features to represent searchers IN better. To help bridge the semantic gap between searchers IN. In particular, we focused on using neurophysiological characteristics within the MIR domain. Furthermore, we examined the EEG recordings of Perceived and Imagined music stimuli to assess their capabilities within a brain-driven MIR system. The findings produced by this study demonstrate the ability of our techniques to achieve Perceived song classification significantly higher than that achieved by prior studies. As well as this, we are the first to report above high accuracy classification of Imagined song stimuli from EEG data.

## References

1. Allegretti, M., Moshfeghi, Y., Hadjigeorgieva, M., Pollick, F.E., Jose, J.M., Pasi, G.: When relevance judgement is happening? an eeg-based study. In: Proceedings of the 38th international acm sigir conference on research and development in information retrieval. pp. 719–722 (2015)
2. Belkin, N.J.: Anomalous states of knowledge as a basis for information retrieval. Canadian journal of information science **5**(1), 133–143 (1980)

3. Belkin, N.J., Oddy, R.N., Brooks, H.M.: Ask for information retrieval: Part ii. results of a design study. Journal of documentation (1982)
4. Belkin, N.J., Oddy, R.N., Brooks, H.M.: Ask for information retrieval: Part i. background and theory. J. Documentation **38**, 61–71 (1997)
5. Chen, J.X., Jiang, D.M., Zhang, Y.N.: A hierarchical bidirectional gru model with attention for eeg-based emotion classification. IEEE Access **7**, 118530–118540 (2019). https://doi.org/10.1109/ACCESS.2019.2936817
6. Chen, X., Ye, Z., Xie, X., Liu, Y., Gao, X., Su, W., Zhu, S., Sun, Y., Zhang, M., Ma, S.: Web search via an efficient and effective brain-machine interface. In: Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. pp. 1569–1572 (2022)
7. Cole, C.: A theory of information need for information retrieval that connects information to knowledge. Journal of the American Society for Information Science and Technology **62**(7), 1216–1231 (2011). https://doi.org/https://doi.org/10.1002/asi.21541, https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21541
8. Cole, C.: A theory of information need for information retrieval that connects information to knowledge. Journal of the American Society for Information Science and Technology **62**(7), 1216–1231 (2011)
9. Davis III, K.M., Kangassalo, L., Spapé, M., Ruotsalo, T.: Brainsourcing: Crowdsourcing recognition tasks via collaborative brain-computer interfacing. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–14 (2020)
10. Eugster, M.J., Ruotsalo, T., Spapé, M.M., Kosunen, I., Barral, O., Ravaja, N., Jacucci, G., Kaski, S.: Predicting term-relevance from brain signals. In: Proceedings of the 37th International ACM SIGIR Conference on Research  Development in Information Retrieval. p. 425–434. SIGIR '14, Association for Computing Machinery, New York, NY, USA (2014). https://doi.org/10.1145/2600428.2609594, https://doi.org/10.1145/2600428.2609594
11. Gao, Z.K., Li, Y.L., Yang, Y.X., Ma, C.: A recurrence network-based convolutional neural network for fatigue driving detection from eeg. Chaos: An Interdisciplinary Journal of Nonlinear Science **29**(11), 113126 (2019)
12. Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., et al.: Meg and eeg data analysis with mne-python. Frontiers in neuroscience p. 267 (2013)
13. Gwizdka, J., Hosseini, R., Cole, M., Wang, S.: Temporal dynamics of eye-tracking and eeg during reading and relevance decisions. Journal of the Association for Information Science and Technology **68**(10), 2299–2312 (2017)
14. Hofmann, T.: Collaborative filtering via gaussian probabilistic latent semantic analysis. In: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval. pp. 259–266 (2003)
15. Hsu, J.L., Zhen, Y.L., Lin, T.C., Chiu, Y.S.: Personalized music emotion recognition using electroencephalography (eeg). In: 2014 IEEE International Symposium on Multimedia. pp. 277–278 (2014). https://doi.org/10.1109/ISM.2014.19
16. Huigen, E., Peper, A., Grimbergen, C.: Investigation into the origin of the noise of surface electrodes. Medical and biological engineering and computing **40**(3), 332–338 (2002)
17. Ingwersen, P., Järvelin, K.: The Turn: Integration of Information Seeking and Retrieval in Context. Springer, Germany (2005)

18. Jacucci, G., Barral, O., Daee, P., Wenzel, M., Serim, B., Ruotsalo, T., Pluchino, P., Freeman, J., Gamberini, L., Kaski, S., et al.: Integrating neurophysiologic relevance feedback in intent modeling for information retrieval. Journal of the Association for Information Science and Technology **70**(9), 917–930 (2019)
19. Jordan, K.G.: Emergency eeg and continuous eeg monitoring in acute ischemic stroke. Journal of clinical neurophysiology **21**(5), 341–352 (2004)
20. Kangassalo, L., Spapé, M., Jacucci, G., Ruotsalo, T.: Why do users issue good queries? neural correlates of term specificity. In: Proceedings of the 42nd international acm sigir conference on research and development in information retrieval. pp. 375–384 (2019)
21. Kangassalo, L., Spapé, M., Ravaja, N., Ruotsalo, T.: Information gain modulates brain activity evoked by reading. Scientific reports **10**(1), 1–10 (2020)
22. Karameh, F.N., Dahleh, M.A.: Automated classification of eeg signals in brain tumor diagnostics. In: Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No. 00CH36334). vol. 6, pp. 4169–4173. IEEE (2000)
23. Kauppi, J.P., Kandemir, M., Saarinen, V.M., Hirvenkari, L., Parkkonen, L., Klami, A., Hari, R., Kaski, S.: Towards brain-activity-controlled information retrieval: Decoding image relevance from meg signals. NeuroImage **112**, 288–298 (2015)
24. Keshavarz, H.: Human information behaviour and design, development and evaluation of information retrieval systems. Program **42**(4), 391–401 (2008)
25. Kingphai, K., Moshfeghi, Y.: On eeg preprocessing role in deep learning effectiveness for mental workload classification. In: International Symposium on Human Mental Workload: Models and Applications. pp. 81–98. Springer (2021)
26. Kuhlthau, C.C.: Inside the search process: Information seeking from the user's perspective. Journal of the American society for information science **42**(5), 361–371 (1991)
27. Maior, H.A., Ramchurn, R., Martindale, S., Cai, M., Wilson, M.L., Benford, S.: fnirs and neurocinematics. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–6 (2019)
28. Michalkova, D., Parra-Rodriguez, M., Moshfeghi, Y.: Information need awareness: An eeg study. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 610–621. SIGIR '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3477495.3531999, https://doi.org/10.1145/3477495.3531999
29. Midha, S., Maior, H.A., Wilson, M.L., Sharples, S.: Measuring mental workload variations in office work tasks using fnirs. International Journal of Human-Computer Studies **147**, 102580 (2021)
30. Morita, Y., Huang, H.H., Kawagoe, K.: Towards music information retrieval driven by eeg signals: Architecture and preliminary experiments. In: 2013 IEEE/ACIS 12th International Conference on Computer and Information Science (ICIS). pp. 213–217 (2013). https://doi.org/10.1109/ICIS.2013.6607843
31. Moshfeghi, Y.: Neurasearch: Neuroscience and information retrieval. In: CEUR Workshop Proceedings. vol. 2950, pp. 193–194 (2021)
32. Moshfeghi, Y., Pinto, L.R., Pollick, F.E., Jose, J.M.: Understanding relevance: An fmri study. In: Advances in Information Retrieval: 35th European Conference on IR Research, ECIR 2013, Moscow, Russia, March 24-27, 2013. Proceedings 35. pp. 14–25. Springer (2013)
33. Moshfeghi, Y., Pollick, F.E.: Search process as transitions between neural states. In: Proceedings of the 2018 World Wide Web Conference. pp. 1683–1692 (2018)

34. Moshfeghi, Y., Pollick, F.E.: Neuropsychological model of the realization of information need. Journal of the Association for Information Science and Technology **70**(9), 954–967 (2019)
35. Moshfeghi, Y., Triantafillou, P., Pollick, F.: Towards predicting a realisation of an information need based on brain signals. In: The world wide web conference. pp. 1300–1309 (2019)
36. Moshfeghi, Y., Triantafillou, P., Pollick, F.E.: Understanding information need: An fmri study. In: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 335–344. SIGIR '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2911451.2911534, https://doi.org/10.1145/2911451.2911534
37. Moshfeghi, Y., Triantafillou, P., Pollick, F.E.: Understanding information need: An fmri study. In: Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. pp. 335–344 (2016)
38. Ntalampiras, S., Potamitis, I.: A statistical inference framework for understanding music-related brain activity. IEEE Journal of Selected Topics in Signal Processing **13**(2), 275–284 (2019). https://doi.org/10.1109/JSTSP.2019.2905431
39. Paisalnan, S., Moshfeghi, Y., Pollick, F.: Neural correlates of realisation of satisfaction in a successful search process. Proceedings of the Association for Information Science and Technology **58**(1), 282–291 (2021)
40. Paisalnan, S., Pollick, F., Moshfeghi, Y.: Towards understanding neuroscience of realisation of information need in light of relevance and satisfaction judgement. In: Machine Learning, Optimization, and Data Science: 7th International Conference, LOD 2021, Grasmere, UK, October 4–8, 2021, Revised Selected Papers, Part I. pp. 41–56. Springer (2022)
41. Pang, S., Hu, X., Cai, Z., Gong, J., Zhang, M.: Building change detection from bi-temporal dense-matching point clouds and aerial images. Sensors **18**(4), 966 (2018)
42. Ras, Z.W., Wieczorkowska, A.: Advances in music information retrieval, vol. 274. Springer (2010)
43. van Rijsbergen, C.J.: (invited paper) a new theoretical framework for information retrieval. In: Proceedings of the 9th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 194–200. SIGIR '86, Association for Computing Machinery, New York, NY, USA (1986). https://doi.org/10.1145/253168.253208, https://doi.org/10.1145/253168.253208
44. Savolainen, R.: Information need as trigger and driver of information seeking: a conceptual analysis. Aslib Journal of Information Management (2017)
45. Schaefer, R.S., Farquhar, J., Blokland, Y., Sadakata, M., Desain, P.: Name that tune: Decoding music from the listening brain. NeuroImage **56**(2), 843–849 (2011). https://doi.org/https://doi.org/10.1016/j.neuroimage.2010.05.084, https://www.sciencedirect.com/science/article/pii/S1053811910008402, multivariate Decoding and Brain Reading
46. Schedl, M., Gómez, E., Urbano, J., et al.: Music information retrieval: Recent developments and applications. Foundations and Trends® in Information Retrieval **8**(2-3), 127–261 (2014)
47. Sonawane, D., Miyapuram, K.P., Rs, B., Lomas, D.J.: Guessthemusic: song identification from electroencephalography response. In: Proceedings of the 3rd ACM India Joint International Conference on Data Science & Management of Data (8th ACM IKDD CODS & 26th COMAD). pp. 154–162 (2021)

48. Spape, M., Davis, K., Kangassalo, L., Ravaja, N., Sovijarvi-Spape, Z., Ruotsalo, T.: Brain-computer interface for generating personally attractive images. IEEE Transactions on Affective Computing **1**(1) (2021)
49. Stober, S.: Toward studying music cognition with information retrieval techniques: Lessons learned from the openmiir initiative. Frontiers in psychology **8**, 1255 (2017)
50. Stober, S.: Toward studying music cognition with information retrieval techniques: Lessons learned from the openmiir initiative. Frontiers in Psychology **8** (2017). https://doi.org/10.3389/fpsyg.2017.01255, https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01255
51. Stober, S., Sternin, A., Owen, A.M., Grahn, J.A.: Towards music imagery information retrieval: Introducing the openmiir dataset of eeg recordings from music perception and imagination. In: International Society for Music Information Retrieval Conference (2015)
52. Taylor, R.S.: The process of asking questions. American documentation **13**(4), 391–396 (1962)
53. Taylor, R.S.: Question-negotiation an information-seeking in libraries. Tech. rep., LEHIGH UNIV BETHLEHEM PA CENTER FOR INFORMATION SCIENCE (1967)
54. Taylor, R.S.: Question-negotiation and information seeking in libraries. Coll. Res. Libr. **76**, 251–267 (1968)
55. de la Torre-Ortiz, C., Spapé, M.M., Kangassalo, L., Ruotsalo, T.: Brain relevance feedback for interactive image generation. In: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. pp. 1060–1070 (2020)
56. Tzallas, A.T., Tsipouras, M.G., Fotiadis, D.I.: Epileptic seizure detection in eegs using time–frequency analysis. IEEE transactions on information technology in biomedicine **13**(5), 703–710 (2009)
57. Van der Veen, V., Dutt-Sharma, N., Cavallaro, L., Bos, H.: Memory errors: The past, the present, and the future. In: Research in Attacks, Intrusions, and Defenses: 15th International Symposium, RAID 2012, Amsterdam, The Netherlands, September 12-14, 2012. Proceedings 15. pp. 86–106. Springer (2012)
58. Wissbrock, F.: Information need assessment in information retrieval; beyond lists and queries. In: Proceedings of the 27th German Conference on Artificial Intelligence (2004)
59. Ye, Z., Xie, X., Liu, Y., Wang, Z., Chen, X., Zhang, M., Ma, S.: Brain topography adaptive network for satisfaction modeling in interactive information access system. In: Proceedings of the 30th ACM International Conference on Multimedia. pp. 90–100 (2022)
60. Ye, Z., Xie, X., Liu, Y., Wang, Z., Chen, X., Zhang, M., Ma, S.: Towards a better understanding of human reading comprehension with brain signals. In: Proceedings of the ACM Web Conference 2022. pp. 380–391 (2022)
61. Zhang, J., Liu, Y., Mao, J., Xie, X., Zhang, M., Ma, S., Tian, Q.: Global or local: Constructing personalized click models for web search. In: Proceedings of the ACM Web Conference 2022. pp. 213–223 (2022)
62. Zhao, R., Yan, R., Wang, J., Mao, K.: Learning to monitor machine health with convolutional bi-directional lstm networks. Sensors **17**(2), 273 (2017)