Predicting the Perception of Tinnitus based on Daily Life Data of the TrackYourTinnitus mHealth Platform in Combination with Country Origin and Season

# Johannes Allgaier1, a, Winfried Schlee2, b, Thomas Probst3, d, and Ru¨ diger Pryss1, e

1Institute of Clinical Epidemiology and Biometry, University of Wu¨ rzburg, Germany

2Department for Psychiatry and Psychotherapy, University of Regensburg, Germany

3Department for Psychotherapy and Biopsychosocial Health, Danube University Krems, Austria

[ajohannes.allgaier@uni-wuerzburg.de](mailto:ajohannes.allgaier@uni-wuerzburg.de) [bwinfr](mailto:bwinfried.schlee@gmail.com)[ied.schlee@gmail.com](mailto:ied.schlee@gmail.com)

[dthomas](mailto:dthomas.probst@donau-uni.ac.at)[.probst@donau-uni.ac.at](mailto:.probst@donau-uni.ac.at)

[eruediger.pryss@uni-wuerzburg.de](mailto:eruediger.pryss@uni-wuerzburg.de)

# ABSTRACT

Tinnitus is a phantom auditory perception in the absence of external sound stimulations. Most people with tinnitus report severe constraints in their daily life. As tinnitus is characterized by a heterogeneity of the patient’s symptoms, researchers often combine tinnitus data with other data sources to reveal new insights, also denoted as multi-modal data fusion. In the context of differences across countries and seasons based on mobile health data, so far, less insights have been presented. Therefore, data of the TrackYourTinnitus mHealth platform (TYT) has been investigated to get a better understanding of season-related differences in the symptom profiles of TYT users as well as to deal with differences based on the country origin of TYT users. Our conducted analyzes address two research questions. First, with respect to provided daily answers of TYT users, in combination with information about the season, the country origin as well as demographic data, we analyzed whether the momentary tinnitus can be related to these aspects. Therefore, a gradient boosting machine from the field of machine learning was trained and analyzed with the goal to predict the momentary tinnitus. Based on the mentioned features, we are able to predict the momentary tinnitus with an accuracy of 93.3%, indicating differences in tinnitus of TYT users with respect to the season and country origin. Second, we analyzed country- and season-specific differences in the light of different perspectives based on descriptive statistics. For example, it could be revealed that tinnitus varies with the temperature in certain countries. The considered perspectives, in turn, have been derived through the TYT data set and its possibilities. The presented results show that the season and the country origin seem to be valuable features when being combined with longitudinal mHealth data of tinnitus patients.

# Introduction

Tinnitus is widely known as a long-term noise in the ears, which is described by patients through heterogeneous sound manifestations[1](#_bookmark9). Economically, tinnitus induces a high burden, as about 10 - 15% of the worldwide population[2](#_bookmark10),[3](#_bookmark11) is affected by this chronic disorder. 2.4% of these affected patients severely suffer from tinnitus day by day[4](#_bookmark12), while 1 to 2 % experience a reduction in their quality of life due to tinnitus, including insomnia, anxiety, hearing difficulties, or depression[5](#_bookmark13)–[7](#_bookmark14). At present, no general treatment, which is able to effectively reduce tinnitus loudness and related fluctuation, exists, which complicates the situation for many patients. In addition, the revealed heterogeneity of tinnitus patients’ symptoms[8](#_bookmark15),[9](#_bookmark16) is challenging for researchers that aim to develop general treatment methods. However, on an individual basis, tinnitus can be reduced, for example, by the use of cognitive behavioral therapies[10](#_bookmark17).

Against the discussed tinnitus characteristics, various efforts are constantly made to learn more about the heterogeneity of symptom profiles of tinnitus patients. However, data sources are often missing to investigate aspects with respect to this heterogeneity of symptom profiles that seem to be interesting. As the proliferation of smartphones has led to powerful mobile health solutions (denoted as mHealth solutions) that are able to establish data sources with opportunities to better deal with differences of symptom profiles. In this paper, such mHealth data source is investigated for tinnitus patients. Although respective investigations have gained attention recently, many opportunities are still not utilized. For example, a comparison of mHealth data of tinnitus users across countries does not exist to the best of our knowledge. In addition, detailed insights based on season differences are also not considered in the context of collected mHealth tinnitus data so far. Therefore, these two questions on differences across seasons and the country origin have been selected for further investigations on symptom profiles of tinnitus

patients using a mHealth platform.

In the context of the mentioned differences, only little research has been presented. In addition, these works are all beyond the scope of mHealth. There is one study on seasonal changes in tinnitus symptomatology, which concludes that searches for tinnitus aspects are higher in winter than in summer in some countries[11](#_bookmark18). Another work suggests an association of depression and season. It provides Internet-based evidence for the epidemiology of seasonal depression. The results suggest that Internet searches for depression by people at higher latitudes are more affected by seasonal changes, while this phenomenon is faded out in tropical areas[12](#_bookmark19). However, already more than 70 years ago, it was clinically observed that tinnitus increases during the winter months[13](#_bookmark20),[14](#_bookmark21). Seasonal affective disorders (SAD), in turn, were studied by the authors of[15](#_bookmark22). They conclude that SADs are present when a symptom occurs during the winter months and disappear completely in summer.

When aiming at mHealth solutions to investigate these differences, at first, the type of collected data must be taken into account as mHealth solutions can be based on different methods, strategies and concepts. In this work, Ecological Momentary Assessments (EMAs) shall be the basis for the investigations as they are particularly appropriate for the investigations at hand[16](#_bookmark23). However, EMA only defines the strategy how participants of a study (usually, longitudinal studies) will be questioned. Three aspects are the main pillars of the EMA strategy: EMAs must be carried out in real life (opposed to a clinical environment) and at arbitrary points in time (to capture the moment of a participant). Finally, a concrete measurement (e.g., though a questionnaire) must be accomplished. If EMAs are now performed through the boundaries of a year and across countries, the data source that can be established through such measurements poses a powerful basis to investigate country- and season-specific differences. Recall that EMA only defines the strategy. In the context of mHealth, digital phenotyping techniques express an important trend to use smartphones to practically enable Ecological Momentary Assessments (EMAs). As smartphones are present in daily life of almost anyone, the performance of EMAs through smartphones can effectively capture the daily life of users over time. Respective evaluations based on mHealth data, in turn, have been recognized as potential alleys for a better support of patients[17](#_bookmark24). MHealth apps, in turn, are the major instrument to operationalize digital phenotyping and EMAs. Many mHealth apps have been presented in this context[18](#_bookmark25)–[20](#_bookmark26). Although data sources have been established by the use of digital phenotyping pose potential, mHealth data comes also with drawbacks[21](#_bookmark27). In the context of tinnitus, the TrackYourTinnitus platform (TYT), which is based on mobile crowdsensing techniques[22](#_bookmark28) as well as EMAs[23](#_bookmark29), puts digital phenotyping into practice. TYT was initially developed to investigate questions about the aforementioned heterogeneity of symptom profiles of tinnitus patients[20](#_bookmark26),[24](#_bookmark30),[25](#_bookmark31). Apart from TYT, other mHealth apps have been proposed to address questions on tinnitus[10](#_bookmark17),[26](#_bookmark32),[27](#_bookmark33), which indicates that digital phenotyping is also promising in the context of tinnitus.

The mentioned investigations on differences across seasons and the country origin have been identified to be possible on the TYT data source. For the concrete analyzes, we have decided to work on the following two major research questions (RQ):

* RQ1: Is the momentary tinnitus predictable by knowledge about the country, season, age, sex, and answers from the daily EMA questionnaires of TYT users (i.e., the EMA questionnaire)?
* RQ2: Are we able to reveal country- and season-specific differences for the reported momentary tinnitus based on the daily EMA questionnaires of TYT users?

Regarding RQ1, we will present results of a machine learning analysis. As TYT was able to gather more than 100,000 EMA questionnaires since 2013 comprised of many dimensions, we decided to answer RQ1 based on machine learning algorithms. We already revealed interesting results on TYT EMA-data based on machine learning[28](#_bookmark34) and its use has been generally recognized in the context of mHealth data in the last years with much attention and valuable results[29](#_bookmark36)–[32](#_bookmark37).

Regarding RQ2, we will present descriptive statistics about the identified country- and season-specific differences. We have detailed the research question into four sub-questions due to the following reason. Based on the two main goals to investigate country- and season-specific (which represent the two categories of differences), we were able to derive two further promising questions. The first one (RQ23) is a combined perspective of the country and the season, while the second one (RQ24) is inspired by medical experts. The following list presents the four sub-questions:

i RQ21: Are there country-specific differences for the momentary tinnitus? ii RQ22: Are there season-specific differences for the momentary tinnitus?

1. RQ23: In the light of a combination of country- and season-specific differences, the question arose, whether the momentary tinnitus varies within the year and across countries.
2. RQ24: The question arose, whether country- and season-specific differences of the reported worst symptom can be identified.

Two additional notes are important regarding RQ21-RQ24. First, the last sub-question was made as mentioned due to the involved medical experts as severe symptoms play an important role in the context of tinnitus research. As TYT asks about nine possible worst symptoms, we investigated how the worst symptom differs across countries and seasons. As the combined perspective taken for RQ23 was useful, this combined perspective was also accomplished for RQ24. Second, in the context of season-specific differences, we added an additional dimension, the temperature course throughout the year, which is inspired by the results of[12](#_bookmark19).

# 1 Results

In this section, results for the research questions are presented subsequently. All research questions are focused on the first question one of the daily TYT questionnaire *Did you perceive the tinnitus right now?*. We refer to this question as the momentary tinnitus in the following.

## RQ1: Is the momentary tinnitus predictable by knowledge about the country, season, age, sex, and an- swers from the daily questionnaire of TYT users (i.e., the EMA questionnaire)?

The machine learning task at hand is a binary classification task. We wanted to know whether it is possible to predict the occurrence of tinnitus for an individual of the TYT platform.

**The features.** We used four different groups of features. The first group of features are dummy features indicating whether an individual comes from that country or not. As 111 countries would lead to an unnecessary increase in the size of the features, we only took those 10 countries with the most filled out daily questionnaires. These countries are [’DE’, ’US’, ’NL’, ’CH’, ’GB’, ’CA’, ’RU’, ’AT’, ’IT’, ’NO’]. The second group of features are the four seasons, which are also coded as dummy features. The third group contains age and sex.

Thomas: Brauchen wir da Referenzen, dass Alter und Geschlecht hinzugefügt worden sind, da bei Tinnitus aussagekräftig? The age is represented by the year of birth. Sex contains two unique values, male and female. The last group of features is a subset of questions of the daily questionnaire. This subset contains information about the momentary mood, arousal, stress level, and concentration. This results in a dataframe with 20 features, 1 binary target, and 74,360 samples from 2,179 users.

**Data preparation.** From previous research works, we already knew that the dataset is imbalanced regarding the target. This means that about 75,000 answers are tinnitus = yes, but only 20,000 tinnitus = no. A classifier that guessed randomly the outcome would get 50 % accuracy on average, a *naive* classifier would simply always predict *Tinnitus=yes* and would get

78.95 % accuracy on average. We therefore draw randomly 54,566 times a sample from the *Tinnitus=no* group, add it to the dataframe, and finally shuffle the samples. This forces each naive classifier to an accuracy down to 50 %. This, in turn, means that any improvement in accuracy can be attributed to the learning of the classifier.

**Machine Learning.** We used a Gradient Boosting Machine[33](#_bookmark38), which builds an additive model and learns subsequently from prior classification trees. We further divided the data into three sets: Two for cross-validation (the training and the validation sets), and one for the final testing. For cross-validation, we used 70 %, for testing, and 30% for validation. We stratified on y while splitting in order to retain the 50-50 distribution of the binary target. After an hyperparameter tuning using gridsearch, we got a **final accuracy of 93.3 %** in the testing set. Details are given in Fig. [1](#_bookmark0).

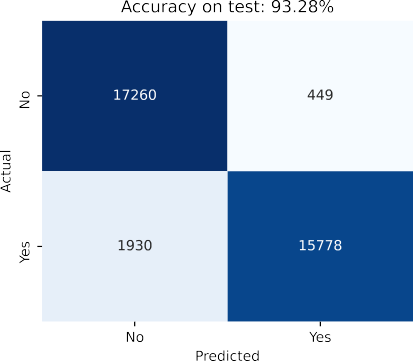
**Feature Importance** To find out which of the variables have a high impact on tinnitus prediction, we looked at the feature importance of the gradient boosting machine. In order to determine the feature importance more accurately, we have investigated three methods for this. The first one is called *gini importance*, the second one is *permutation importance*, while the last one is the *correlation*. The different methods measure the feature importance in different units. To make the results comparable, we have also created an importance ranking.

If we first divide the features into their groups (country, season, demographics, EMA), we can see that the EMA features (questions 4, 5, 6, and 7) and the demographic features (sex, age) seem to be the most important feature groups on average. The third most important feature group is the season, followed by the countries. Year of birth is a very important feature for the Gradient Boosting Machine for two reasons. First, it has a high cardinality (many different values) and second, it has a moderate correlation with current tinnitus. The permutation importance of 29.7 % suggests that the accuracy becomes 29.7

% percentage points worse when the year of birth is replaced by a random variable. For example, almost all Russian users have consistently answered the question about current tinnitus in the affirmative. Within the countries feature, Russia therefore has a high correlation with current tinnitus. However, because there are relatively few users compared to all users, the gini importance for RU only shows a value of 2.19 %.

Thomas: Sollte man zwischen den RQs Überleitungssätze bauen?

|  |  |  |  |
| --- | --- | --- | --- |
| **precision** | **recall** | **F1 score** | **support** |
| **Tinnitus No** 0.90 | 0.97 | 0.94 | 17709 |
| **Tinnitus Yes** 0.97 | 0.89 | 0.93 | 17708 |
| **accuracy** |  | 0.93 | 35417 |
| **macro avg** 0.94 | 0.93 | 0.93 | 35417 |
| **weighted avg** 0.94 | 0.93 | 0.93 | 35417 |

1. Confusion Matrix for the gradient boosting classifier. Although there is a little tendency on false negatives, the overall accuracy of 93.3 % on the test set is significantly better than random guessing.



1. Classification report for the gradient boosting classifier. The tendency to predict no rather than yes leads to a larger F1 score for Tinnitus=No and a larger recall for Tinnitus=Yes.

**Figure 1.** Confusion matrix and classification report for the gradient boosting machine used to predict whether an individual has momentary tinnitus or not.

**Importances Ranks**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Features** | **gini** | **permutation** | **correlation** | **gini** | **permutation** | **correlation** |
| **AT** | 0.35% | 0.19% | 2.88% | 20 | 20 | 14 |
| **CA** | 0.59% | 0.85% | 4.37% | 19 | 16 | 11 |
| **CH** | 1.65% | 1.26% | 9.38% | 14 | 13 | 5 |
| **DE** | 2.37% | 1.99% | 3.64% | 9 | 9 | 13 |
| **GB** | 0.92% | 0.78% | 0.00% | 16 | 17 | 20 |
| **IT** | 0.62% | 0.35% | 7.61% | 18 | 19 | 7 |
| **NL** | 0.93% | 0.99% | 0.25% | 15 | 15 | 17 |
| **NO** | 0.84% | 0.38% | 7.51% | 17 | 18 | 8 |
| **RU** | 2.19% | 1.09% | 13.41% | 10 | 14 | 3 |
| **US** | 2.10% | 2.26% | 7.77% | 11 | 7 | 6 |
| **autumn** | 2.48% | 1.88% | 3.87% | 7 | 10 | 12 |
| **spring** | 1.93% | 1.29% | 0.14% | 12 | 12 | 18 |
| **summer** | 1.86% | 1.66% | 1.59% | 13 | 11 | 16 |
| **winter** | 2.46% | 2.23% | 5.15% | 8 | 8 | 10 |
| **Male** | 3.81% | 4.38% | 6.43% | 6 | 6 | 9 |
| **year\_of\_birth** | 24.89% | 29.77% | -11.57% | 1 | 1 | 4 |
| **question4** | 9.12% | 11.41% | -18.38% | 4 | 4 | 1 |
| **question5** | 6.70% | 8.25% | 0.07% | 5 | 5 | 19 |
| **question6** | 17.45% | 16.46% | 17.84% | 2 | 3 | 2 |
| **question7** | 16.73% | 17.33% | -2.31% | 3 2 | | 15 |

**Table 1.** Feature importances of the gradient boosting machine and correlations (Pointbiserial and Corrected Cramer’s V) of univariate features with tinnitus occurrence. The gini importances add up to 100 %, the permutation importances indicates the increase of the error rate if that features was left out. As year\_of\_birth is a feature with high cardinality, it clearly helps the tree-based gradient boosting machine to determine tinnitus occurrence. However, the correlation is moderately negative.

## RQ2: Are we able to reveal country- and season-specific differences for the reported momentary tinnitus based on the daily questionnaire of TYT users?

To answer this question, there are 97,742 responses from 3,691 users from a total of 111 countries for the period from April 2014 to February 2021. For the further analysis, we restricted ourselves to the countries represented by more than 30 users with more than 300 questionnaires in total. For this subset, with 15 countries, 3,163 users remain with a total of 88,049 filled out daily questionnaires. Most responses are from Germany with 51,804 completed questionnaires, generated by 1,410 users, whereas the fewest completed questionnaires come from the Federative Republic of Brazil with 334 completed questionnaires, generated by 50 users. The mean number of filled out questionnaires per country is 5870 (std = 13,058). The mean number of users is 210 (std = 357). For the question of interest *Did you perceive the tinnitus right now?* (question1), mean for ’Yes’ is

78.97 % (std = 12.21 %), an interquartile range of 15.73 %, with a maximum value of 95.58 % from Italy, and a minimum

value of 48.66 % from Norway, was found.

**RQ21: Are there country-specific differences for the momentary tinnitus?** A chi-square test of independence showed that there are significant differences between the countries, *χ*2(14*, N* = 85933) = 2441*.*44*, p < .*001. 105 post-hoc *χ*2 tests were performed to compare pairwise differences. Using corrected p values, 91 pairs of countries were rejected (p = .05). 14 pairs could not be rejected at p = .05, i.e., the pair Germany-Great Britain, and Germany-Sweden. This indicates that these countries have a similar pattern in momentary tinnitus occurrence. A detailed overview of the answers of question1 (*Did you perceive the tinnitus right now?*) is given in Table [2](#_bookmark2).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Country\_Name** |  | **No** | **Yes** | **n\_questionnaires** | **n\_users** |
| Australia, Commonwealth of |  | 14.5% | 85.5% | 666 | 77 |
| Austria, Republic of |  | 29.6% | 70.4% | 1321 | 68 |
| Belgium, Kingdom of |  | 28.6% | 71.4% | 972 | 44 |
| Brazil, Federative Republic of |  | 8.7% | 91.3% | 344 | 50 |
| Canada |  | 13.9% | 86.1% | 2341 | 126 |
| France, French Republic |  | 16.6% | 83.4% | 467 | 72 |
| Germany, Federal Republic of |  | 21.0% | 79.0% | 51804 | 1410 |
| Italy, Italian Republic |  | 4.4% | 95.6% | 1220 | 81 |
| Netherlands, Kingdom of the |  | 33.1% | 66.9% | 7268 | 180 |
| Norway, Kingdom of |  | 51.3% | 48.7% | 1178 | 42 |
| Spain, Kingdom of |  | 9.3% | 90.7% | 517 | 82 |
| Sweden, Kingdom of |  | 18.2% | 81.8% | 362 | 38 |
| Switzerland, Swiss Confederation |  | 32.8% | 67.2% | 5139 | 122 |
| United Kingdom |  | 20.5% | 79.5% | 3713 | 210 |
| United States of America |  | 12.8% | 87.2% | 10737 | 561 |

**Table 2.** Tinnitus Occurrence by country for individuals of the TYT platform grouped by country. When filling out a questionnaire, most users state that they perceive the tinnitus at that moment. The chance for this is 78 %, with a standard deviation of 12 percent.

**RQ22: Are there season-specific differences for the momentary tinnitus?** To answer this question, we again analysed only countries represented by more than 30 users with more than 300 completed questionnaires *per season*. This filter setting holds True for Switzerland, Germany, the United States, Great Britain, and the Netherlands. The largest sample is again for Germany, with 51,534 completed questionnaires, the smallest sample is for the UK, with 3,684 completed questionnaires.

If we do not group by country, the greatest probability for momentary tinnitus is in summer with 83.4% (std = 8.6%). In contrast, the lowest probability for momentary tinnitus is in winter, with 71.0 % (11.8 %). The interquartile range is 14.5 % for winter, and 11.8 % for summer. If we group by country, the highest probability for momentary tinnitus is in summer in Great Britain (95.7 %), the lowest in winter in Switzerland (60.7 %). The ratios of yes-no-responses are shown in Fig. [2](#_bookmark3). Considering not only these five countries, but all 111 countries in the present data set without setting a questionnaires or user threshold, the probability of momentary tinnitus perception is 80.2 % in summer, 80.1 % in fall, 78.5 % in spring, and 75.4

% in winter. A *χ*2 test of independence showed that there was a significant association between season and momentary tinnitus,

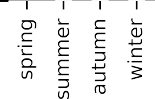
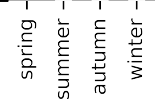
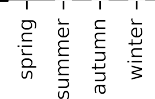
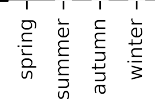
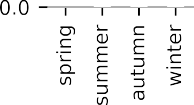
*χ*2(3*, N* = 95446) = 216*.*19*, p < .*001. Overall user reporting for tinnitus is thus most likely in summer.

In a slightly different approach, we considered months instead of seasons. Therefore, we increased the granularity of the x-axis. In addition, we examined the respective average temperature per month in relation to tinnitus occurrence for the countries considered (Switzerland, Germany, U.S., Great Britain, and the Netherlands). When multiple temperature data points from different cities were available for a country, they were aggregated with the average.

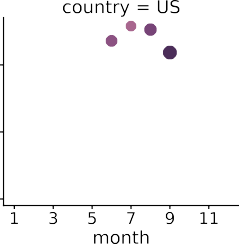
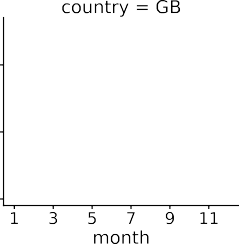
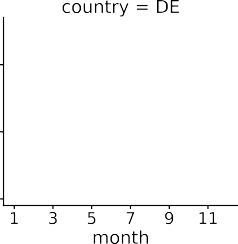
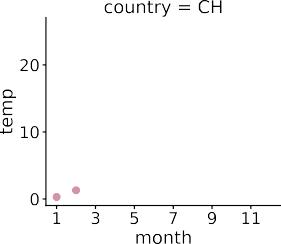
A high positive correlation can be obtained for the Netherlands (r(10) = .83, p < .001), for Great Britain (r(10 = .86, p < .001), and for Switzerland (r(10) = .72, p = .009). On the contrary, the U.S. show a non-significant medium negative correlation (r(10) = -.41, p = .18). For Germany, however, the correlation between temperature and tinnitus occurrence can be considered uncorrelated (r(10) = -.09, p = .78). The cyclical temperature pattern associated with tinnitus over the year for the various countries is shown in Fig. [4](#_bookmark4). There was a statistically significant difference between the countries as determined by one-way ANOVA (F(4, 55)= 6.69 , p < .001). A post-hoc Tukey test indicates that the annual course of momentary tinnitus is different between the country pairs Netherlands-U.S. (p < .01) and Switzerland-U.S. (p < .01).

**RQ23: In the light of a combination of country- and season-specific differences, the question arose, whether the momentary tinnitus varies within the year and across countries.** In contrast to the previous section, we have ignored temperature in this question. Instead, we examined the following: For each of the countries considered, and for each individual month of the year, we calculated the probability of tinnitus by dividing the number of yes responses by the sum of responses. In the following step, we examined the probability of tinnitus over the course of the year. To increase comparability, we





**Figure 2.** Distribution of the tinnitus occurrence (*Did you perceive the tinnitus right now?*) by country and season for Switzerland (CH), Germany (DE), the United States of America (US), the United Kingdom of Great Britain & Northern Ireland (GB), and the Netherlands. *n* denotes the number of filled out daily questionnaires per country for all seasons.

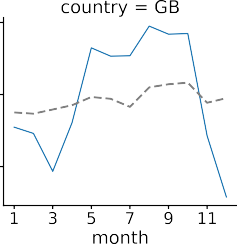


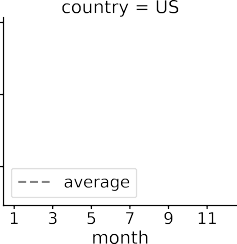
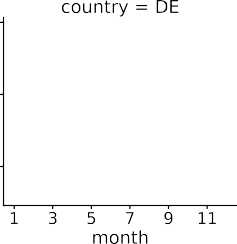
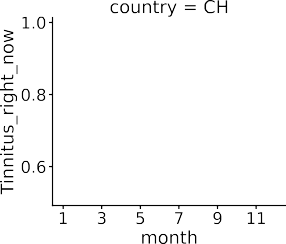
**Figure 3.** Cyclical temperature pattern associated with tinnitus for Switzerland (CH), Germany (DE), the United States of America (US), the United Kingdom of Great Britain & Northern Ireland (GB), and the Netherlands. The x-axis shows the month, the y-axis the temperature in degrees Celsius. The larger the circle, the higher the average probability for a momentary tinnitus for this country in this month. The size and color of the cycles indicate the chance of momentary tinnitus. The bigger the cycle, the higher the chance.

additionally calculated the average of the tinnitus probability for all available data on a monthly basis.

Since most of the data comes from Germany, this country has a correspondingly large influence on the average values. Accordingly, the curve for Germany is very similar to the curve for all data (statistic = .17, p = 1.00). On the contrary, the Netherlands, the U.S. and Switzerland reveal a different distribution of the tinnitus with p-values < 0.01. For Great Britain, the distribution can be considered to be slightly different as p-value is .10. An overview of the distributions compared with with the average is given in Fig. [4](#_bookmark4). A summarizing statistical overview, in turn, is given in Table [3](#_bookmark5).

The highest probability for tinnitus is in America with an average chance of 87 %, the lowest probability in Switzerland with 68 %. The largest variance occurs in Great Britain, with 16 % standard deviation, the smallest in Germany, with 4 %. For this data set, tinnitus occurred least in Switzerland in March (53 %), and most in the UK in August (98 %).





**Figure 4.** Course of occurrence of tinnitus over the year for Switzerland (CH), Germany (DE), the United States of America (US), the United Kingdom of Great Britain & Northern Ireland (GB), and the Netherlands. The x-axis shows the month, the y-axis the probability for tinnitus occurrence. The dashed grey lines show the average of tinnitus occurrence for all data *except* the country plotted on this axis. The graph indicates that people of different nations perceive tinnitus differently throughout the year.

## country count mean std min 25% 50% 75% max

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **CH** | 12.00 | 0.68 | 0.08 0.53 0.64 | 0.67 | 0.74 | 0.79 |
| **DE** | 12.00 | 0.79 | 0.04 0.73 0.76 | 0.78 | 0.82 | 0.87 |
| **GB** | 12.00 | 0.80 | 0.16 0.52 0.69 | 0.81 | 0.94 | 0.99 |
| **NL** | 12.00 | 0.71 | 0.13 0.59 0.61 | 0.66 | 0.76 | 0.95 |
| **US** | 12.00 | 0.87 | 0.05 0.76 0.85 | 0.89 | 0.91 | 0.94 |

**Table 3.** Statistics for the occurrence of tinnitus throughout the year grouped by country. For this data set, momentary tinnitus occurred least in Switzerland in March (53 %), and most in the UK in August (98 %).

**RQ24: The question arose, whether country- and season-specific differences of the reported worst symptom can be identified.** To answer this research question, we again focused on the five countries [CH, DE, GB, NL, US]. When registering on the TYT platform, the question about the worst tinnitus symptom is asked once. For each country and season, we calculated the relative number of answers within a country to compare which symptom is more likely in which season. Each column adds up to 100 %. The 1,310 users from Germany have the lowest standard deviation (.94 std). The Netherlands with 175 users has the largest standard deviation (2.01 std). *I find it harder to relax* is the most likely symptom in the Netherlands in fall, with 8.57 %, and, at the same time, with a global maximum. *Feeling depressed* ranks second for the UK and the Netherlands. For the U.S., the two worst symptoms are *difficulty following a movie or conversation* and *concentration problems*. For the U.S., however, there is little variation between seasons within these two worst symptoms. *None of these symptoms* ranks second for Switzerland. *Irritability with friends and family* is the least indicated worst symptom for all countries.

In a similar approach, we disregarded countries and investigated the evolution of the worst tinnitus symptom between seasons. Thus, we examined whether there are different worst symptoms per season. *Because of the tinnitus I am more irritable with my family, friends and colleagues* is the most unlikely symptom (mean = 5.9 %, std = 1.0 %. The most likely symptom constitutes *I find it harder to relax because of the tinnitus* (mean = 17.7 %, std = 1.9 %). Details are given in Fig. [5](#_bookmark6). Difficulties in relaxing is the worst symptom across all seasons. The data further indicates that feelings of depression are stronger in the months of autumn and winter. Difficulties in following conversations are more pronounced in summer. Irritability with colleagues or family is the least selected symptom. However, a chi-square test of independence showed that there was no significant association between worst symptom and season, *χ*2(24, N = 4) = 0, p = NS.

# Discussion

The present work investigated the differences of momentary tinnitus in relation to seasons and temperatures. Although we found significant differences between seasons and countries, this does not establish causality between the variables. Although our findings potentially provide important insights for further tinnitus research, there are a few limitations that should be discussed. First, there may be myriad of other reasons why tinnitus is more likely in some countries in summer and in some in winter. Influencing factors could be, for example, air pressure, the stress level or the number of hours of sunshine. Second, user numbers vary widely between countries. This can lead to a selection bias in the evaluation. That means, if one user was particularly active in filling out the daily questionnaire, and the other 29+x users were not, this might lead to a selection bias. Third, although our research results indicate different seasonal trends for tinnitus for different countries, there may be

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **worst\_symptom** | **season** | **CH DE**  **(n=114) (n=1310)** | | **GB**  **(n=201)** | **NL**  **(n=175)** | **US**  **(n=537)** |
| **Because of the tinnitus I am more**  **irritable with my family, friends**  **and colleagues.** | **spring**  **summer autumn winter** | 0.00%  0.00% | 1.91%  1.37% | 0.50%  1.00% | 1.14%  1.14% | 0.93%  1.86% |
| 1.75%  0.88% | 1.68%  1.53% | 0.50%  1.00% | 2.29%  0.57% | 2.23%  0.93% |
| **Because of the tinnitus I am more** | **spring** | 4.39% | 2.67% | 1.99% | 1.14% | 1.49% |
| **sensitive to environmental noises.** | **summer** | 0.88% | 1.83% | 1.00% | 1.71% | 2.23% |
|  | **autumn** | 4.39% | 2.90% | 0.50% | 4.00% | 2.61% |
|  | **winter** | 2.63% | 2.14% | 1.00% | 0.00% | 1.68% |
| **Because of the tinnitus it is** | **spring** | 0.88% | 3.44% | 2.49% | 2.86% | 4.66% |
| **difficult to concentrate.** | **summer** | 2.63% | 2.90% | 1.99% | 1.14% | 2.79% |
|  | **autumn** | 0.88% | 3.66% | 1.49% | 6.29% | 5.21% |
|  | **winter** | 1.75% | 2.60% | 1.00% | 3.43% | 2.79% |
| **Because of the tinnitus it is** | **spring** | 2.63% | 3.36% | 3.48% | 1.14% | 4.28% |
| **difficult to follow a conversation,** | **summer** | 2.63% | 2.98% | 5.97% | 3.43% | 4.66% |
| **a piece of music or a film.** | **autumn** | 3.51% | 4.12% | 2.49% | 3.43% | 3.17% |
|  | **winter** | 2.63% | 3.59% | 2.49% | 1.71% | 3.91% |
| **Because of the tinnitus it is hard** | **spring** | 4.39% | 2.60% | 2.99% | 1.14% | 3.54% |
| **for me to get to sleep.** | **summer** | 1.75% | 1.98% | 2.99% | 0.57% | 2.98% |
|  | **autumn** | 0.88% | 3.36% | 4.98% | 5.14% | 3.17% |
|  | **winter** | 3.51% | 2.67% | 5.47% | 1.71% | 4.10% |
| **I am feeling depressed because of** | **spring** | 3.51% | 1.91% | 2.99% | 3.43% | 0.93% |
| **the tinnitus.** | **summer** | 0.88% | 2.14% | 4.48% | 2.86% | 2.05% |
|  | **autumn** | 4.39% | 2.14% | 4.48% | 5.14% | 3.91% |
|  | **winter** | 1.75% | 1.60% | 6.47% | 2.86% | 2.79% |
| **I don't have any of these** | **spring** | 6.14% | 2.67% | 1.00% | 0.00% | 1.68% |
| **symptoms.** | **summer** | 1.75% | 2.37% | 1.00% | 1.71% | 1.68% |
|  | **autumn** | 4.39% | 3.05% | 1.00% | 2.29% | 2.42% |
|  | **winter** | 7.89% | 2.37% | 1.99% | 2.29% | 2.42% |
| **I find it harder to relax because of** | **spring** | 4.39% | 5.19% | 6.97% | 5.71% | 4.47% |
| **the tinnitus.** | **summer** | 7.89% | 3.44% | 2.49% | 4.57% | 3.54% |
|  | **autumn** | 3.51% | 5.57% | 4.48% | 8.57% | 3.91% |
|  | **winter** | 3.51% | 3.28% | 7.96% | 2.86% | 2.23% |
| **I have strong worries because of** | **spring** | 3.51% | 2.21% | 2.99% | 0.00% | 2.79% |
| **the tinnitus.** | **summer** | 0.88% | 2.60% | 1.49% | 4.57% | 1.49% |
|  | **autumn** | 1.75% | 3.59% | 3.48% | 6.29% | 1.68% |
|  | **winter** | 0.88% | 2.60% | 1.49% | 2.86% | 2.79% |

**Table 4.** Distribution of the worst symptom for each country and season. We only considered countries with more than 300 questionnaires from more than 30 users. Each column adds up to 100 %. *n* denotes the number of users from this country.

individuals who perceive tinnitus seasonally quite differently, possibly even completely in the opposite direction. This means that these findings are not applicable to individuals.

Thomas: Sollte im vorangegangenen Abschnitt mehr belegt werden?

Thomas: Zweite Frage, ich finde es eher skeptisch zu schreiben, wie das Johannes gemacht hat besser ,

aber sollten wir positiver werden?

For the worst tinnitus symptom per country and season, comparability between countries and seasons may also be biased by the selection due to the low number of users per category. For Switzerland, for example, we would expect 3.17 individuals per symptom per season (i.e., 2.8 % per line), if symptoms and seasons were equally distributed. In this respect, it is surprising for Switzerland, for example, that relaxation is more difficult in summer (7.89 %) than in winter (3.51 %). The situation is different with Germany. Here, we have a large number of users of 1310 and would expect 36.4 individuals per category, if the symptoms were equally distributed among all seasons. This argument is supported by the fact that the variance in Germany is lower than in Switzerland. Nevertheless, we can observe for Germany that relaxing due to tinnitus is more difficult for spring and autumn (about 5 % ) than for summer or winter (about 3 %).

**Feature Importance** High cardinality features such as year\_of\_birth and the daily questions are assigned a higher importance as these features can be easily split up into multiple, potentially pure subsets. For binary features, the tree classifier can only split up the data once. However, for features with high cardinality, the tree can potentially split up the data n\_unique

- 1 times. Feature importance does not establish causality between input variables and target. It is rather an estimator of which

**Figure 5.** Development of the worst symptom for tinnitus over the seasons. Users are asked this question once when completing the baseline questionnaire (n = 3458). Irritability with friends and family is the least selected symptom, difficulty with relaxation is the most selected. Difficulty following a conversation has a clear high in summer. The values for each season add up to 100 %.



variable has the greatest predictive power for the gradient boosting machine. Any other classifier, such as a neural network, would potentially produce a different ranking for feature importance.

**Worst Season for Tinnitus** This question cannot be answered unambiguously and conclusively. Related work on tinnitus and seasonality does suggest winter as the worst season. However, 41.8 % of individuals (n = 100) report perceiving summer as the second worst season, which argues against the theory of seasonal affective disorders. In the study, which aggregated tinnitus search requests from online platforms by season and country, winter was also highlighted as a more frequent season. In the study, which aggregated tinnitus search requests from online platforms by season and country, winter was also highlighted as a more frequent season[11](#_bookmark18). However, the results are different even for countries with similar longitudes. For example, this is the case for Sweden and the United Kingdom. The noise in the results could be due to confounder variables or the mentioned selection bias.

# Materials and Methods

The study was approved by the Ethics Committee of the University Clinic of Regensburg (ethical approval No. 15-101-0204). All users read and approved the informed consent before participating in the study. The study was carried out in accordance with relevant guidelines and regulations.

**The questionnaires.** For the tinnitus prediction task, three linked data sets were used. The first one refers to the baseline questionnaire named named *Tinnitus Sample Case History Questionnaire (TSCHQ)*. This questionnaire is completed by each

TYT user *once* when starting the app for the fist time. In this questionnaire, demographic data as well as data about the individual course of the tinnitus are collected, such as the onset of the tinnitus or the worst symptom that is related to tinnitus. Baseline characteristics from this questionnaire for the five countries CH, DE, GB, NL, US, as well as all other countries, can be seen in Table [6](#_bookmark7). For the characteristics *handedness* and *family history of tinnitus complaints*, a *χ*2 test was performed. The *χ*2 test showed that there was no significant association within the country groups, *χ*2(8, N=2319) = 6.64, p=0.58 for *handedness*, and *χ*2(4, N=2314) = 4.33, p=0.36, for *family history*. To compare the age distributions between the countries, a one-way ANOVA was performed with F(4, 2267) = 5.17, p < 0.001. A post-hoc pairwise Tukey test revealed differences between DE and US (meandiff = 2.36, p < 0.05), and GB and US (meandiff = 5.07, p < 0.01). The remaining 8 pairwise groups have no significant differences in their means.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Age**  **F(4, 2267) = 5.17, p < 0.001** | | | | | | | | **Handesness**  **X²(8, N=2319) = 6.64, p=0.58** | | | **Family History**  **X²(4, N=2314) = 4.33, p=0.36** | |
| **Country** | **Sex** | **Count** | **Mean** | **Std** | **Min** | **25%** | **50%** | **75%** | **Max** | **Left** | **Both Sides** | **Right** | **No** | **Yes** |
| **CH** | **Female** | 32 | 48.38 | 13.84 | 31 | 37 | 47 | 62 | 74 | 0.0% | 9.1% | 90.9% | 69.7% | 30.3% |
|  | **Male** | 78 | 49.94 | 13.95 | 21 | 39 | 50 | 59 | 78 | 12.5% | 17.5% | 70.0% | 71.3% | 28.8% |
| **DE** | **Female** | 414 | 44.36 | 13.80 | 8 | 33 | 46 | 55 | 79 | 10.5% | 13.1% | 76.4% | 74.3% | 25.7% |
|  | **Male** | 851 | 49.15 | 13.83 | 10 | 39 | 50 | 58 | 87 | 10.6% | 13.1% | 76.3% | 79.3% | 20.7% |
| **GB** | **Female** | 91 | 41.81 | 12.33 | 17 | 32 | 42 | 51 | 70 | 8.8% | 15.4% | 75.8% | 74.7% | 25.3% |
|  | **Male** | 106 | 46.12 | 13.13 | 13 | 37 | 46 | 57 | 71 | 13.2% | 7.5% | 79.2% | 78.5% | 21.5% |
| **NL** | **Female** | 25 | 50.76 | 12.07 | 29 | 43 | 47 | 61 | 73 | 5.9% | 14.7% | 79.4% | 73.5% | 26.5% |
|  | **Male** | 95 | 45.79 | 14.12 | 18 | 34 | 50 | 57 | 73 | 14.0% | 8.1% | 77.9% | 73.5% | 26.5% |
| **US** | **Female** | 242 | 47.71 | 13.19 | 12 | 38 | 49 | 57 | 84 | 14.9% | 8.9% | 76.2% | 69.6% | 30.4% |
|  | **Male** | 284 | 51.58 | 12.68 | 16 | 43 | 54 | 60 | 81 | 11.5% | 12.8% | 75.7% | 78.9% | 21.1% |
| **all** | **Female** | 1102 | 44.46 | 13.60 | 8 | 33 | 45 | 55 | 84 | 11.2% | 13.4% | 75.3% | 72.7% | 27.3% |
|  | **Male** | 2231 | 47.15 | 13.95 | 1 | 37 | 48 | 57 | 114 | 12.9% | 15.7% | 71.4% | 78.2% | 21.8% |

**Table 5.** Statistical comparison of the five countries CH, DE, GB, NL, and US with the rest of all users. Additionally, the data is grouped by gender. For the *χ*2 tests, the N differs from the Count column, as some data is missing. The *χ*2 for *handedness* and *family history* is not significant. For the comparison of the age distributions, the post-hoc Tukey test shows significant mean differences for Germany with the United States (p < 0.05), and Great Britain with the United States (p < 0.01).

When logging onto the TYT platform, users are asked for their worst tinnitus symptom. This symptom can be one of the following.

* I am feeling depressed because of the tinnitus.
* I find it harder to relax because of the tinnitus.
* I have strong worries because of the tinnitus.
* Because of the tinnitus it is difficult to follow a conver- sation, a piece of music or a film.
* Because of the tinnitus it is hard for me to get to sleep.
* Because of the tinnitus it is difficult to concentrate.
* Because of the tinnitus I am more irritable with my family, friends and colleagues.
* Because of the tinnitus I am more sensitive to environ- mental noises.
* I don’t have any of these symptoms.

As we also record fill-in dates of answers to this questionnaire, and the country of the user, we can link the worst symptom to both the season and country. To assign the fill-in date to a season, we used the astronomical seasons as a guide. More specifically, spring starts in March 21st, summer in June 21st, autumn in September 23rd, and winter in December 21st.

The second data set refers to the *daily questionnaire*. It includes eight questions about the current tinnitus state, i.e., the tinnitus situation and the feelings of the individual *right now*. However, the eighth *dynamic* question depends on the worst symptom of the individual from the TSCHQ questionnaire and asks whether the individual has this specific worst symptom right now or not. If an individual answered *I don’t have any of these symptoms* in the beginning, no eighth question appears in the daily questionnaire. Consequently, the number of answers for question 8 depends on the number of individuals that have selected this worst symptom in questionnaire TSCHQ. On the other hand, the number of answers for questions one to seven equals each other. These questions are seen by every individual in the same way and are as follows:

1. Did you perceive the tinnitus right now?
2. How loud is the tinnitus right now?
3. How stressful is the tinnitus right now?
4. How is your mood right now?
5. How is your arousal right now?
6. Do you feel stressed right now?
7. How much did you concentrate on the things you are doing right now?
8. *This question depends on the worst symptom selected in the questionnaire TSCHQ.*

Depending on the features that are selected for the classification task, the number of examples *m* depends on the dynamic question eight. The questions for *mood* and *arousal* are questions using a self-assessment scale (SAM)[34](#_bookmark39), with 9 possible values. Depending on a user’s operating system, the answer is stored with different accuracy. Therefore, rounding errors can occur in the hundredths range on Android phones. We neglected these rounding errors in pre-processing considering the amount of 18 other features.

## Data preprocessing

The raw data comes from three .csv files, which, in turn, are extractions from the TYT database. The first file is a dataframe containing meta information from all registered users (number of users = 8685 by Feb. 2021). This meta data includes among others the country, nationality, and mobile platform. The second file is the baseline questionnaire and contains 3700 users that filled out the initial questionnaire. The daily questionnaire is the last file with 3044 users that answered 98,074 daily questionnaires. We can see from this that of the registered users, about one in three completes the daily questionnaire at least once.

The user\_id is mandatory to merge the three data sets. As a consequence, all rows where user\_id equals NULL, we dropped that row. We further removed the 25 test-users with known user ids to reduce bias and noise in the data. The remaining merged dataframe has 97,742 rows and 65 columns. This dataframe has been used for for the statistical analyses provided in the results section.

**Machine Learning Preprocessing** For the machine learning task, a further preprocessing is required. Gradient boosting machines can only handle numerical data with no missing values. We therefore dropped rows that contained missing values, which affected about 24 % of the data. We then needed to convert categorical features into numbers. As decision trees split data in binary groups, we used the pandas.get\_dummies() method to convert the countries and seasons into several columns. The column name is then the category. A 1 indicates, that this category applies, i.e., autumn = 1, which, in turn, means that the other seasons must be zero. In order not to increase the number of columns unnecessarily, we used the drop\_first = True keyword argument. This means, we get k-1 dummies out of k categorical levels by removing the first level. The last step considered the imbalanced distribution of the target variable *tinnitus occurrence*. About 79 % of the users reported *yes*. Any naive machine learning classifier would therefore simply always predict *yes*, regardless of the input of features and would still get 79 % accuracy on average. Using the F1 accuracy score, the performance can be measured better, but the classifier would still be over-trained on positive examples. We therefore bootstrapped negative examples with replacement until we had a balanced dataset. The final dataset has 118,054 samples with 22 features each.

**Estimation of features importances.** The values of Table [1](#_bookmark1) were calculated using three different methods, the gini impor- tance, the permutation importance, and the correlation metric. Depending on the feature scaling, two different correlation metrics have been applied. If the input feature is categorical, Corrected Cramer’s V[35](#_bookmark40) was applied. If it is continuous, the Point Biserial method[36](#_bookmark41) was used. Cramer’s V is defined in range (0, 1), whereas the Point Biserial correlation is defined in range (-1, 1). Nevertheless, to be able to order the results *within* the column, we took the absolute value from the Point Biserial result. Although all results are in percentages, it is not possible to compare them line by line. This is due to the different units of measurement. Therefore, we have created the ranking. For the gini and the permutation importances, both methods are used using the trained gradient boosting machine. The gini importances is an impurity based method. The higher it is, the more important the feature. Notably, within this column, all values add up to 100 %. The importance of a feature is calculated as the reduction of the impurity caused by this feature. For the permutation importance, the percentage values are an estimate for the increase of the error rate on average if that features would have been replace by a random feature. That means, if the variable gender would be replaced with a random variable, the error would increase by 6.43 %-points. That column does not neccessarily add up to 100 %.

## Gradient Boosting Classifier for binary classification

Why did we chose the Gradient Boosting machine? It is a tree-based Machine Learning algorithm and related to Random Forests. Machine Learning contests on the Kaggle platform have recently shown that this algorithm is superior to most Deep Learning methods when it comes to tabular data such as house pricing prediction problems. Both, Random Forests and Gradient Boosting Machines use several trees to predict the outcome. However, one of the main differences between those two algorithms is the *time aspect*. That is, the Gradient Boosting algorithm learns from previous miss-classified samples by putting more weight on those. Furthermore, it does not easily tend to overfitting like decision trees do.

We used the Python implementation from scikit-learn[37](#_bookmark42) to apply the Gradient Boosting machine to the dataset. We then defined the 20 features (10 countries, 4 seasons, sex, year of birth, mood, arousal, stress, concentration level) and the target

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **variable\_name** | **variable meaning** | **mean** | **std** | **scaling** |
| **AT** | Austria | 0.02 | 0.13 |  |
| **CA** | Canada | 0.03 | 0.16 |  |
| **CH** | Switzerland | 0.08 | 0.27 |  |
| **DE** | Germany | 0.62 | 0.49 |  |
| **GB** | Great Britain | 0.05 | 0.21 |  |
| **IT** | Italy | 0.01 | 0.10 |  |
| **NL** | Netherlands | 0.07 | 0.25 |  |
| **NO** | Norway | 0.02 | 0.13 | binary |
| **RU** | Russia | 0.02 | 0.14 |  |
| **US** | United States | 0.09 | 0.29 |  |
| **spring** |  | 0.26 | 0.44 |  |
| **summer** | season | 0.24 | 0.43 |  |
| **autumn** |  | 0.25 | 0.43 |  |
| **winter** |  | 0.25 | 0.44 |  |
| **Male** | Sex | 0.74 | 0.44 |  |
| **year\_of\_birth** | Year of birth | 1967.85 | 12.68 | integer |
| **question4** | How is your mood right  now?  How is your arousal right  now? | 0.58  0.25 | 0.20  0.22 | SAM from 0 to 1 with stepsize 0.125 |
| **question5** |
| **question6** | Do you feel stressed right  now? How much did you concentrate on the things  you are doing right now? | 0.26  0.59 | 0.23  0.31 | Slider in range (0, 1) |
| **question7** |
| **question1** | Did you perceive the  tinnitus right now? | 0.50 | 0.50 | binary |

**Table 6.** Overview of the features and the target used to train the gradient boosting classifier. Most of the features are binary, the year of birth has the highest cardinality. The whole dataset has shape (118054, 22).

(tinnitus occurrence right now). The whole dataset was divided into three sets: Training, development, and testing. Training + development got 70 % of the data, testing 30 %. To avoid a selection bias, we stratified on y. Setting a random\_state (also known as seed) ensured that the results are reproducible. For the tuning of the hyperparameters, we used a gridsearch approach. Within that, we varied the learning\_rate, the max\_depth of each tree, the sizes of the subsamples, the minimum number of samples per leaf, and the fraction of randomly chosen features per tree. 1,280 combinations of the hyperparameters have been evaluated systematically, the final chosen setup can be seen in Listing [1](#_bookmark8). Each combination was cross-validated within the development set using a 5-fold split. This means that the 70 % of the training and validation data was divided into 5 folds. Four of each were used for training and one for validation. The **best mean test score** on validation was 91.1 %, with a standard deviation of .002. On the test dataset, an accuracy of 93.3 % was achieved.

When leaving out the features sex and age, the mean test score dropped to 88.9 % using the same hyperparameters. Using only the binary features seasons and countries leads to a decrease of the accuracy on the test set down to 58 %. This is caused by the low dimensional feature space.

1 # Gridsearch setup

2 params\_gb = { ’ l e a r n i n g \_ r a t e ’ : [ 0 . 1 , 0 . 2 , 0 . 3 , 0 . 5 , 1 ] ,

3 ’ max\_depth ’ : [ 3 , 4 , 5 , 10 ] ,

1. ’ verbose ’ : [ 1 ] ,
2. ’ random\_state ’ : [ 4 2 ] ,

6 ’ subsample ’ : [ 0 . 2 5 , 0 . 5 , 0 . 75 , 1 ] ,

7 ’ min\_samples\_leaf ’ : [ 1 , 2 , 3 , 10 ] ,

8 ’ max\_features ’ : [ 0 . 2 5 , . 5 , . 75 , 1 ]

9 }

10

1. # Chosen hyperparameters
2. G r a d i e n t B o o s t i n g C l a s s i f i e r ( l os s = ’ deviance ’ , l e a r n i n g \_ r a t e = 0 . 5 , n\_estimators =100 , subsample = 1 . 0 , c r i t e r i o n = ’ friedman\_mse ’ , min\_samples\_split =2 , min\_samples\_leaf =1 , m i n \_ w e i g h t \_ f r a c t i o n \_ l e a f = 0 . 0 , max\_depth=10 ,

min\_impurity\_decrease = 0 . 0 , m i n \_ i m p u r i t y \_ s p l i t =None , i n i t =None , random\_state =42 , max\_features = 0 . 5 , verbose =0 ,

max\_leaf\_nodes=None , warm\_start=False , v a l i d a t i o n \_ f r a c t i o n = 0 . 1 , n\_iter\_no\_change=None , t o l =0. 0001 , ccp\_alpha

= 0 . 0 )

13

14 \ l a b e l { l i s t i n g }

**Listing 1.** Hyperparameter set up for the Gradient boosting machine classifier

# References

1. Kiang, N., Moxon, E. & Levine, R. Auditory-nerve activity in cats with normal and abnormal cochleas. *Sensorineural* *hearing loss* 241–273 (1970).
2. Davis, A. & Rafaie, E. A. Epidemiology of tinnitus. *Tinnitus handbook* **1**, 23 (2000).
3. Langguth, B. A review of tinnitus symptoms beyond ‘ringing in the ears’: a call to action. *Curr. medical research opinion*

**27**, 1635–1643 (2011).

1. Halford, J. B. & Anderson, S. D. Anxiety and depression in tinnitus sufferers. *J. psychosomatic research* **35**, 383–390 (1991).
2. Langguth, B., Kreuzer, P. M., Kleinjung, T. & De Ridder, D. Tinnitus: causes and clinical management. *The Lancet Neurol.*

**12**, 920–930 (2013).

1. Izuhara, K. *et al.* Association between tinnitus and sleep disorders in the general japanese population. *Annals Otol. Rhinol.* *& Laryngol.* **122**, 701–706 (2013).
2. McKENNA, L., HALLAM, R. S. & HINCHCLIFFEf, R. The prevalence of psychological disturbance in neuro-otology outpatients. *Clin. Otolaryngol. & Allied Sci.* **16**, 452–456 (1991).
3. Cederroth, C. R. *et al.* Towards an understanding of tinnitus heterogeneity. *Front. aging neuroscience* **11**, 53 (2019).
4. Cederroth, C. R. *et al.* Medicine in the fourth dimension. *Cell metabolism* **30**, 238–250 (2019).
5. Mehdi, M. *et al.* Contemporary and systematic review of smartphone apps for tinnitus management and treatment. (2020).
6. Plante, D. T. & Ingram, D. G. Seasonal trends in tinnitus symptomatology: evidence from internet search engine query data. *Eur. Arch. Oto-Rhino-Laryngology* **272**, 2807–2813 (2015).
7. Yang, A. C., Huang, N. E., Peng, C.-K. & Tsai, S.-J. Do seasons have an influence on the incidence of depression? the use of an internet search engine query data as a proxy of human affect. *PloS one* **5**, e13728 (2010).
8. Hilger, J. A. Autonomic dysfunction in the inner ear. *The Laryngoscope* **59**, 1–11 (1949).
9. Atkinson, M. Tinnitus aurium: some considerations concerning its origin and treatment. *Arch. otolaryngology* **45**, 68–76 (1947).
10. Miller, A. L. Epidemiology, etiology, and natural treatment of seasonal affective disorder. *Altern. medicine review* **10**

(2005).

1. Shiffman, S., Stone, A. A. & Hufford, M. R. Ecological momentary assessment. *Annu. Rev. Clin. Psychol.* **4**, 1–32 (2008).
2. Unnikrishnan, V. *et al.* The effect of non-personalised tips on the continued use of self-monitoring mhealth applications.

*Brain Sci.* **10**, 924 (2020).

1. Torous, J., Friedman, R. & Keshavan, M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. *JMIR mHealth uHealth* **2**, e2 (2014).
2. Martínez-Pérez, B., De La Torre-Díez, I. & López-Coronado, M. Mobile health applications for the most prevalent conditions by the world health organization: review and analysis. *J. medical Internet research* **15**, e120 (2013).
3. Schlee, W. *et al.* Momentary assessment of tinnitus—how smart mobile applications advance our understanding of tinnitus. In *Digital Phenotyping and Mobile Sensing*, 209–220 (Springer, 2019).
4. Rowland, S. P., Fitzgerald, J. E., Holme, T., Powell, J. & McGregor, A. What is the clinical value of mhealth for patients?

*NPJ Digit. Medicine* **3**, 1–6 (2020).

1. Pryss, R. Mobile crowdsensing in healthcare scenarios: taxonomy, conceptual pillars, smart mobile crowdsensing services. In *Digital Phenotyping and Mobile Sensing*, 221–234 (Springer, 2019).
2. Kraft, R. *et al.* Combining mobile crowdsensing and ecological momentary assessments in the healthcare domain. *Front.* *Neurosci.* **14**, 164 (2020).
3. Schlee, W. *et al.* Measuring the moment-to-moment variability of tinnitus: the trackyourtinnitus smart phone app. *Front.* *aging neuroscience* **8**, 294 (2016).
4. Probst, T., Pryss, R., Langguth, B. & Schlee, W. Emotional states as mediators between tinnitus loudness and tinnitus distress in daily life: Results from the “trackyourtinnitus” application. *Sci. reports* **6**, 1–8 (2016).
5. Sereda, M., Smith, S., Newton, K. & Stockdale, D. Mobile apps for management of tinnitus: users’ survey, quality assessment, and content analysis. *JMIR mHealth uHealth* **7**, e10353 (2019).
6. Mehdi, M. *et al.* Smartphone apps in the context of tinnitus: Systematic review. *Sensors* **20**, 1725 (2020).
7. Unnikrishnan, V. *et al.* Predicting the health condition of mhealth app users with large differences in the number of recorded observations-where to learn from? In *International Conference on Discovery Science*, 659–673 (Springer, 2020).
8. Aguilera, A. *et al.* mhealth app using machine learning to increase physical activity in diabetes and depression: clinical trial protocol for the diamante study. *BMJ open* **10**, e034723 (2020).
9. Said, A. B., Mohamed, A., Elfouly, T., Abualsaud, K. & Harras, K. Deep learning and low rank dictionary model for mhealth data classification. In *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 358–363 (IEEE, 2018).
10. Qureshi, K. N., Din, S., Jeon, G. & Piccialli, F. An accurate and dynamic predictive model for a smart m-health system using machine learning. *Inf. Sci.* **538**, 486–502 (2020).
11. Cheung, Y. K. *et al.* Are nomothetic or ideographic approaches superior in predicting daily exercise behaviors? analyzing n-of-1 mhealth data. *Methods information medicine* **56**, 452 (2017).
12. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals statistics* 1189–1232 (2001).
13. Pryss, R. *et al.* Prospective crowdsensing versus retrospective ratings of tinnitus variability and tinnitus–stress associations based on the trackyourtinnitus mobile platform. *Int. J. Data Sci. Anal.* **8**, 327–338 (2019).
14. Bergsma, W. A bias-correction for cramér’s v and tschuprow’s t. *J. Korean Stat. Soc.* **42**, 323–328 (2013).
15. Tate, R. F. Correlation between a discrete and a continuous variable. point-biserial correlation. *The Annals mathematical* *statistics* **25**, 603–607 (1954).
16. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
17. Schlee, W. *et al.* Innovations in doctoral training and research on tinnitus: The european school on interdisciplinary tinnitus research (esit) perspective. *Front. aging neuroscience* **9**, 447 (2018).

# Acknowledgements

This work was partly funded by the ESIT (European School for Interdisciplinary Tinnitus Research[38](#_bookmark43)) project, which is financed by European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 722046 and the UNITI (Unification of Treatments and Interventions for Tinnitus Patients) project financed by the European Union’s Horizon 2020 Research and Innovation Programme, Grant Agreement Number 848261.

# Author contributions statement

J.A. primarily wrote this paper, created the figures, tables, and trained the machine learning algorithms for gender prediction. W.S., B.L., T.P. and R.P. carefully read and revised the paper. Everybody contributed to the methodology. R.P. supervised the paper.

# Supplementary Information

The Python code to replicate the Machine Learning classifiers, figures and tables is available on [github](https://github.com/joa24jm/tyt_gender_prediction.git).

# Additional Information

The authors declare no competing interests.