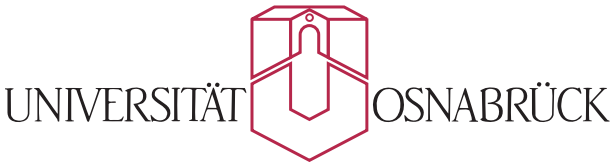

Decoding emotional content from GIFs using MEG data

by

Vera Klütz

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
Department of Cognitive Science
University of Osnabrück

First Supervisor: PD. Dr. rer. nat. Gordon Feld
Second Supervisor: Prof. Dr. rer. nat. Gordon Pipa
Submitted: September 5, 2024

Eidesstattliche Erklärung

Hiermit erkläre ich, Vera Klütz, die vorliegende Arbeit *Decoding emotional content from GIFs using MEG data* selbständig verfasst zu haben und keine anderen Quellen oder Hilfsmittel als die angegebenen verwendet zu haben.

Mannheim, den 5.9.2024

Vera Klütz, 991846

Affirmation Statement

I, Vera Klütz, hereby certify that the work presented here is, to the best of my knowledge and belief, original and the result of my own investigations, except as acknowledged, and has not been submitted, either in part or whole, for a degree at this or any other university.

Mannheim, September 5, 2024

Vera Klütz, 991846

Abstract

1 page This is going to be great! Keep on reading!

Acknowledgements

less than 1 page Thank you everyone!

Contents

List of Figures

List of Tables

1 Introduction

ca 40 pages for whole thesis
- 3-4 pages
- Motivation to do this project,
- background
- aim of the project
- shortly say which methods are being used
- (Structure of this thesis?)

1.1 Section One

Hello. This is a citation: (Engel, Maye, Kurthen, & König, 2013).

1.2 Section Two

Hello. This is a table (ref. Table 1.1).

Table 1.1: LMM Regression Results Overview

Model:	MixedLM	Dependent Variable:	Gaze
No. Observations:	1400	Method:	REML
No. Groups:	10	Scale:	1.0
Min. group size:	90	Log-Likelihood:	-1700.0
Max. group size:	130	Converged:	Yes
Mean group size:	120		

hjkhj gjgh

2 Current State of Research

15 pages

Literature overview and stand der Forschung
theoretical concepts and models
conceptual framework
'Something like: In order to assess the current state of research for this task, it has to be divided into three subtasks. For each of one of them, major choices have to be made in order to successfully

2.1 Measuring Emotions

In order to be able to decode emotional content it has to be stated what 'emotional' means. Since the adjective links to the broader concept of 'emotions', the question arises what emotions are and how to measure them for this task.

Ekman and Cordaro define that "Emotions are discrete, automatic responses to events, forming a family of related states with at least 12 characteristics" (Ekman & Cordaro, 2011). This definition highlights the automatic nature of emotional responses as well as the connectedness between those emotional states. However, it is debatable which characteristics fully describe the different states and how many core emotions there are. Vytal and Hamann conducted a meta-analysis and for five basic emotions they identified consistent neural correlates, namely fear, anger, disgust, sadness, and happiness (Vytal & Hamann, 2010). It is important to note that this meta-analysis does not divide emotions based on philosophical conclusion or by observing facial expressions, but by looking at neural correlates, which is a quantifiable dimension. For this thesis, being able to measure emotions is a necessary premise and neural activity is a numerically measurable physiology. Still, the continuous neural activity has to be mapped to discrete emotions, as Ekman and Cordaro already pointed out above.

Another way to describe an emotion, which is easier to standardize between people and less dependent on the usage of the same linguistic terms, is by using the valence and arousal model. According to Colibazzi et al., "Emotions are linear combinations of valence and arousal" (Colibazzi et al., 2010). This means that an emotion can be assessed by its position on these two scales, with valence ranging from 'negative' to 'positive' and arousal diverging from 'calming' to 'exciting' (Kensinger, 2004).

On the contrary, Cowen and Keltner (2017) oppose to the valence arousal

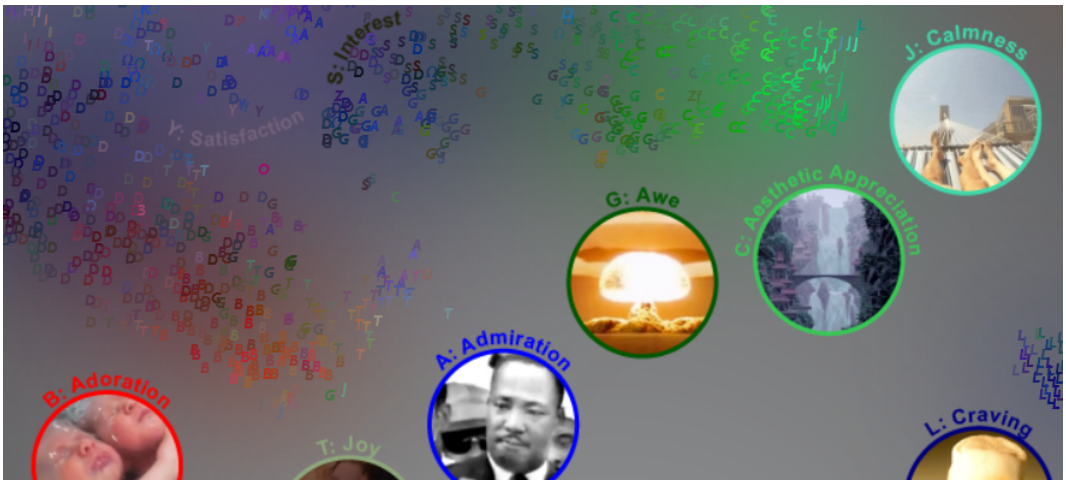

Figure 1: Excerpt from the full chromatic map of 27 distinct emotion categories by Cowen and Keltner (2017). Map accessed on 30.08.2024 on `https://s3-us-west-1.amazonaws.com/emogifs/map.html`

model and suggest to divide emotions into 27 distinct categories, like 'Joy', 'Admiration', and 'Awe'. They note that the boundaries between those categories, however, are rather fuzzy than discrete (Cowen & Keltner, 2017), which can be seen in Figure 1. The map shows datapoints in a colored semantic space, the full chromatic map can be seen in the Appendix in Figure 10. Each data point is a short video that is supposed to elicit an emotion and is assigned to a category by color. In total, 2,185 videos have been mapped, and additionally been rated on 14 affective dimensional scales, including valence and arousal.

Other stimuli have been used to elicit emotions in different studies. These includes pictures (Kim et al., 2018) (Polo et al., 2023) (Uhrig et al., 2016), videos (Uhrig et al., 2016) (Xue, Wang, Hu, Bi, & Lv, 2022) (Gross & Levenson, 1995), audio (Polo et al., 2023), odor, haptic modalities (Gatti, Calzolari, Maggioni, & Obrist, 2018), and a combination of those (Polo et al., 2023) (Xue et al., 2022). It is unclear which modality performs best to elicit emotions, Zhou et al. (2014) state that visual stimuli are as effective as auditory ones and Xue et al. highlight that the combination of odor and video can enhance emotional experience. Even the modalities themselves differ between studies, for example an auditory stimuli can consist of music or spoken stories.

If the chosen stimulus is effective, an emotion will be elicited in the brain. Kragel and LaBar (2016) found that distributed neural systems which span over cortical and subcortical regions are activated in order to uniquely represent affec-

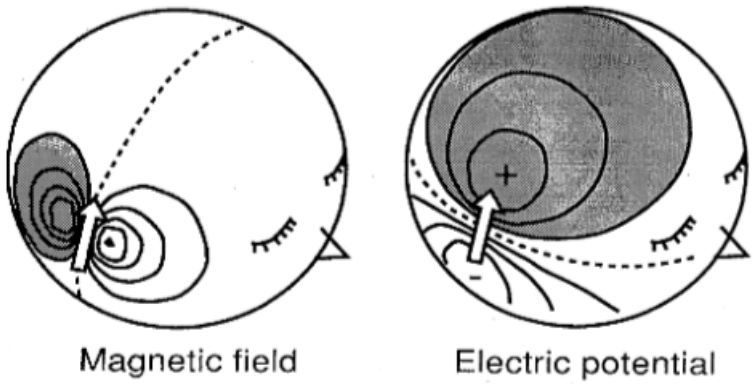

Figure 2: Idealized scheme of a tangential dipole causing an electric potential and a magnetic field (Hämäläinen et al., 1993, p.417)

tive dimensions and emotions. Each neuron that is activated as part of this process will transport and alter an electric current which itself will cause a magnetic field (Trepel, 2012). The numerosity of neurons firing together will produce a dipole that is causing an electric current and therefore also a magnetic field, which can be seen in Figure 2 (Hämäläinen et al., 1993). The resulting signal is measurable from outside the scalp, for example the magnetic field can be measured in the femtotesla range, meaning $\mathbf{10^{-15}}$, so it is a very weak signal (Hämäläinen et al., 1993).

2.2 MEG Data Preprocessing

There are various modalities that can pick up on those very weak signals and therefore are able to measure processes in the brain. The magnetoencephalography (MEG) can measure magnetic fields and the electroencephalography (EEG) can measure the strength of electric potentials (Baillet, Mosher, & Leahy, 2001). Both are relying on the underlying neural activity which is causing the electric and magnetic signal and both have a very good time resolution below 100 ms (Baillet et al., 2001). As a comparison, a magnetic resonance imaging (MRI) has a time resolution of ca. 1 second, but excels with a much better spatial resolution than EEG and MEG (Baillet et al., 2001).

2.2.1 MEG

The MEG is a non-invasive neuroimaging technique that uses an array of gradiometers to reconstruct the magnetic field.

2.2.2 Data Preprocessing

https://www.sciencedirect.com/topics/engineering/neural-signal-processing

2.3 Machine Learning

For the scope of this thesis, several Gigabytes of data have been collected. AI is a useful tool for interpreting such data, especially for detecting common patterns and features which would be difficult to identify for a human. Haenlein and Kaplan (2019) define AI as the ability of a system to accurately interpret outside data, learn from it, and apply that learning to accomplish particular tasks and goals through adaptable change. Whereas first AI attempts relied on explicit representation of knowledge, machine learning (ML), which is a subpart of AI, uses computational methods rather than preset formulas to teach machines to learn from their past experiences (Shaveta, 2023). AI can also generally be described as a system for advanced problem solving (Janiesch, Zschech, & Heinrich, 2021), and ML is a branch of AI that relieves the user of the need to explicit their knowledge into machine-interpretable rules (Janiesch et al., 2021). Instead, ML is able to automatically improve with experience (Mitchell, 1997) and is therefore nowadays seen as the core of AI (Mukhamediev et al., 2022).

Machine Learning itself can be divided into three different subparts, which can be seen in Figure 3, namely Supervised Learning, Unsupervised Learning, and Reinforcement Learning. The first one requires training data that contains not only the input variable, but also the labeled target variable (Janiesch et al., 2021), so the 'solution' to the task the machine learning algorithm is trying to solve. This allows the algorithm to learn and update its own parameters, until it is able to correctly predict unseen data (Janiesch et al., 2021). The second category that can be seen in Figure 3 is called Unsupervised Learning, because here the labels are missing, and therefore an external way of 'supervision' is not included (James et al., 2023). Here, the goal is to find structural information (Janiesch et al., 2021). For example, in the clustering technique the algorithm is looking for common properties in groups of elements, or for the dimensionality reduction technique a high-dimensional space is projected into a lower one (Janiesch et al., 2021). In contrast, in a reinforcement learning system the machine learning model learns how to achieve the goal on its own by maximizing a reward through trial and error (Janiesch et al., 2021). Therefore, its current state has to be described, a goal has to be specified, a list of actions that are permissible have to be provided, as well as the environmental constraints that apply to their outcomes (Janiesch et al., 2021).

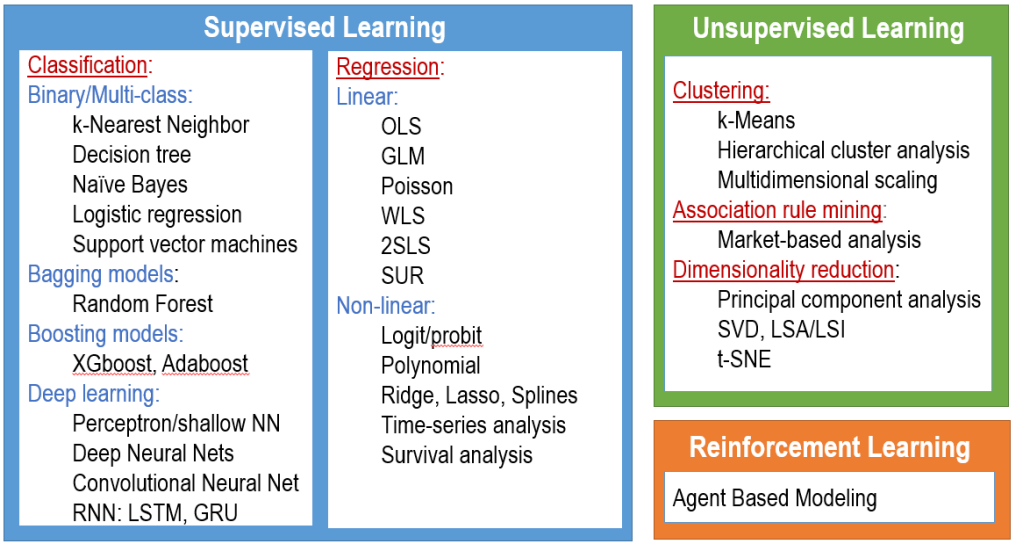

Figure 3: Machine Learning overview (Krumnack, 2022)

For the scope of this thesis, supervised learning is the most relevant. Labeled target variables exist which helps to form clear criteria for model optimization and the labels allow for a straight forward evaluation of the algorithm.

In the field of supervised learning, there are two major fields of problems, namely classification problems and regression problems, as shown on Figure 3. For the first one, categorical values are predicted, for example the category 'happy' or 'sad', for regression numerical values are predicted, for example someones height or age (Janiesch et al., 2021). For each of those two problems there are different classifiers to choose from, as depicted on Figure 3.

2.3.1 Logistic Regression

Despite its name, one common classification method is called 'Logistic Regression'. It estimates class probabilities for binary classes with the logistic function, which ensures that the calculated class probability stays between 0 and 1 (James et al., 2023). The logistic function can be seen in Figure 4 along with the formula for calculating the probability for the target variable Y to belong to a class, given the input variable X.

If more than two classes are part of the paradigm, then multinomial logistic regression can be used to overcome the binary capacities of logistic regression. Therefore, one class is serving as a baseline, which class is chosen does not influence the prediction but only the coefficient estimates of the mathematical model

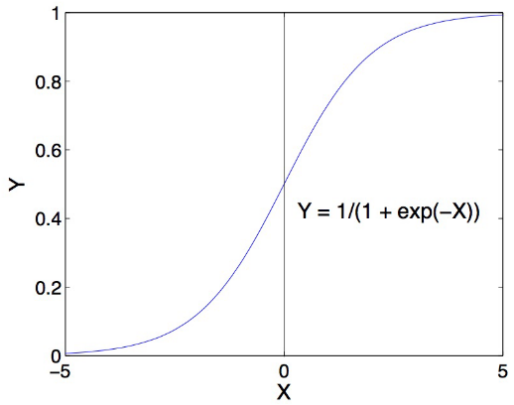

Figure 4: Logistic function (Mitchell, 2020, p.8)

(James et al., 2023). Another way to implement multinomial logistic regression is to use softmax coding. Here, all classes are treated symmetrically instead of choosing a baseline class (James et al., 2023).

2.3.2 Support Vector Machines

Another classifier type which falls into the category of binary/multi-class classifiers as shown in Figure 3 are support vector machines (SVMs). They rely on the notion of a single hyperplane being able to divide the space of data points into two subparts (James et al., 2023). A hyperplane is always one dimension smaller than the actual data space, so in a 2D space a hyperplane is just 1D, and therefore a line (James et al., 2023). In a 3D space it would be a plane dividing the space into two parts. The idea is illustrated in Figure 5.

Of course, the figure just shows one out of many possible hyperplanes for this space, it could be rotated or shifted a bit and would still be correct. The optimal hyperplane position is influenced by all data points that lie on or within a certain margin of the hyperplane, which are called the support vectors (James et al., 2023). However, all of those hyperplanes imaginable for Figure 5 lead to a linear decision boundary (James et al., 2023), but in fact, many classification problems require a non-linear one. Therefore, a support vector machine will want to enlarge the feature space in a way, so that a linear hyperplane can optimally divide the two classes again (James et al., 2023). For more than two classes, the SVM will either compare each of the classes on a 'one class versus another one class' basis or by comparing one class against all other classes that are treated as one big class. This way, the binary SVM classifier is able to deal with multiple classes.

Support vector machines and logistic regression often give quite similar results

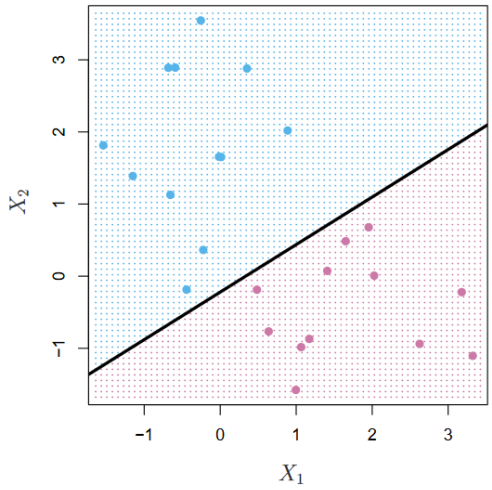

Figure 5: A black hyperplane is visualizing a decision rule that divides a grid into a blue and a red area. Each data point inside this area will be assigned to the corresponding class. (James et al., 2023, p.370)

(James et al., 2023). If the classes are more overlapping, logistic regression is preferred, and for more distinct classes, SVMs perform better (James et al., 2023).

2.3.3 Random Forest

Random Forest is a classifier belonging to the group of so called bagging models as shown in Figure 3. They are based on a decision tree which means that a set of splitting rules used for diving the predictor space can be visualized in a tree (James et al., 2023), which can be seen in Figure 6. The predictor space is the space where all variables lie inside. A decision tree consists of internal nodes, where a splitting rule is applied, and branches which connect the nodes (James et al., 2023). At the bottom of the decision tree are the terminal nodes which represent the leaves of the tree. Here, one segment of the predictor space is assigned to each leave (James et al., 2023). For classification problems, the most commonly occurring class of this segment is assigned to the terminal node, in regression problems it is the mean response that is calculated (James et al., 2023).

The explained procedure will overly adapt to already seen data which has been used to build the tree, but will perform worse for unseen data, because the decision boundaries are too individualized to this particular data set now. This is also referred to as overfitting. In order to avoid this, the decision tree has to be pruned, meaning that some branches and the corresponding nodes will be removed. Ideally, the resulting tree has the best trade off between the two parameters tree com-

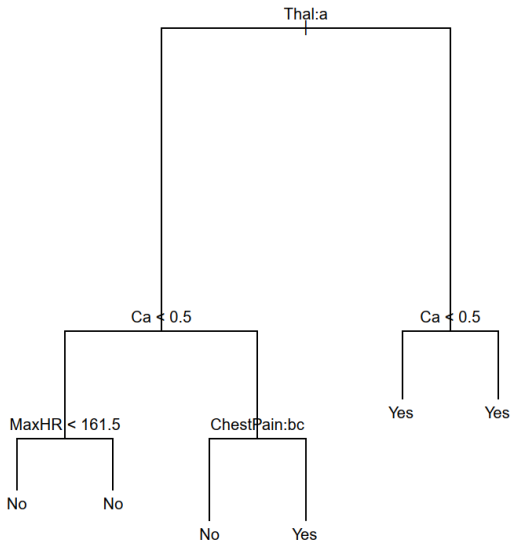

Figure 6: Decision tree for heart disease data, with the two categories Yes and No respectively. The nodes include a Thallium stress test (Thal), Ca values, as well as chest pain and maximal heart rate values. (James et al., 2023, p.340)

plexity and fit of already seen data (James et al., 2023).

Unfortunately, the resulting tree will have a high variance, meaning that if for example the input data would be split in half and two trees would be built, they would differ a lot (James et al., 2023). This is a problem because the resulting decision tree is supposed to be general and to also work the same way, the best way, for unseen data. In order to avoid this, a method called bagging is used. This means, that a number of trees are built based on overlapping subsets of the training data (James et al., 2023). In the end, the trees are averaged to obtain the resulting decision tree.

What is special about the Random Forest approach is that it takes all those trees before averaging and tries to decorrelate them. To do so, when building the tree, at each node only a random subset of predictors are available to be a split candidate (James et al., 2023). This way it can be ensured that a substantial reduction of the variance between the different trees is yielded (James et al., 2023).

2.3.4 Deep Learning

When people nowadays talk about AI, they often actually refer to deep learning applications. The most well known deep learning classifier is probably the neural network, which can be seen as the cornerstone of deep learning (James et al., 2023). A very simple neural network can be seen in Figure 7.

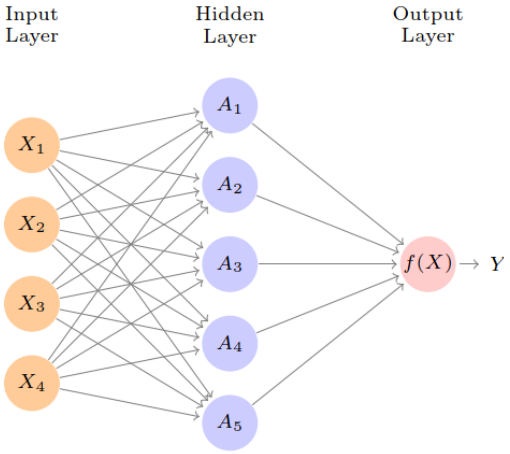

Figure 7: Representation of a neural network. $\vec{X}$ denotes the input vector, $\vec{A}$ represents the activations and $f(\vec{X})$ is the output function. (James et al., 2023, p.400)

Initially, the idea of neural networks was to mimic neurons in the brain (James et al., 2023). The sensory input is represented by the units of the input layer. In the brain, this input gets directed to neurons which are shown as units in the hidden layer. The amount of units in the hidden layer and the amount of hidden layers can be chosen and differ between neural networks (James et al., 2023). In the brain, the neurons get activated and fire if a certain threshold is met, making the information processing process non-linear (James et al., 2023). In neural networks, each unit of the hidden layer has its own weight, and also each unit of the input layer gets weighted. The sum of all weighted input units plus and the respective weight of the hidden unit get passed into a non-linear activation function. This function can for example be a sigmoid function, which looks similar to the one shown in Figure 4. Nowadays, the rectified linear unit activation function (ReLU) is a common choice, as it can be stored and computed more efficiently than the sigmoid activation (James et al., 2023). It equals zero for all negative values and is an identity function otherwise. At the end, the output layer uses the previously calculated activations as input and determines the classification result.

This is just the most simple and basic form of a neural network, deep learning includes much more elaborated and advanced approaches, for example convolutional neural networks for image classification or recurrent neural networks for time series applications (James et al., 2023).

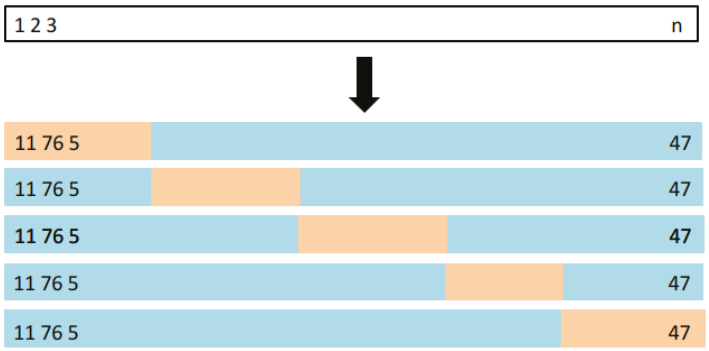

Figure 8: 5-fold cross-validation: the beige parts represent the respective testing data, whereas the blue parts mark the corresponding training data. (James et al., 2023, p.207)

2.3.5 Model performance

In order to build a model, the available data set is usually split into two parts: one for training the model, and one for testing its performance. Common split ratios are 70% or 80% of the data for training and the remaining 30% or 20% are used for testing. This procedure helps to prevent overfitting, which describes a model that has adjusted too well to the training data, including to its noise or coincidental regularities (Mitchell, 1997).

However, especially for small data sets it can happen that the testing data split does not represent a good part of the entire data space. K-fold cross-validation can overcome this issue and yield a more representative test result. Therefore, the dataset is divided into k different parts of roughly equal size, also referred to as folds (James et al., 2023). One fold is taken as testing data, whereas the model is fit to the remaining k-1 folds, which can be seen in Figure 8 (James et al., 2023). In the end, the test error of the entire model is obtained by averaging over the test errors of each of the k models (James et al., 2023).

In order to assess a classification model, different evaluation metrics are available. The most common is accuracy, which tells how much of the testing data has been correctly classified (Wardhani et al., 2019). However, this metric might be misleading in case of imbalanced data. If for example a dataset with 100 datapoints contains 90 positive and only 10 negative examples, the classifier could yield a 90% accuracy by classifying everything as a positive example, without having learned to differentiate between those two classes. A good overview of correctly and incorrectly classified examples gives a confusion matrix, which can be seen in Figure 9.

Actual	Predicted	
	Positive	***Negative***
Positive	TP	FN
Negative	FP	TN

Figure 9: Confusion matrix scheme: F=False T=True P=Positives N=Negatives (Wardhani et al., 2019, p.2)

A confusion matrix consists of four fields, taking into account the predicted values and the actual ones. If the classifier marks an example as positive, and it also actually is one, it is referred to as a true positive (TP) (Wardhani et al., 2019). This goes in line with actual negative examples, which also have been classified as negative, making them true negatives (TN). If a positive has been falsely classified as negative, it becomes a false negative (FN), and if a negative is falsely predicted to be positive, it is a false positive (FP) (Wardhani et al., 2019).

So, more formally expressed, accuracy is the ratio of correctly classified examples divided by all examples and can be described with the formula

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

(Wardhani et al., 2019). Another metric which is better at dealing with imbalanced data is the f1-score. It is build by taking the harmonic mean of precision and recall (Wardhani et al., 2019). Precision describes the ratio of true positives out of all values predicted to be positives, whereas recall describes the amount of true positives in relation to all actually positive values (Wardhani et al., 2019). This leads to the formulas

$$precision = \frac{TP}{TP + FP}$$

(Wardhani et al., 2019)

$$recall = \frac{TP}{TP + FN}$$

(Wardhani et al., 2019)

$$f1 - score = \frac{2}{precision^{-1} + recall^{-1}}$$

(Wardhani et al., 2019)

2.3.6 Regression methods

2.4 Feature Selection

2.5 Related Work

3 Methodology

5-10 pages

Description of Methods:

preprocessing

state/ explain data getting procedure, high arousal, different valences

Machine learning, ´classification vs regression, linear regression

”However, for low-dimensional data input, especially in cases of limited training data availability, shallow ML can still produce superior results (Zhang and Ling 2018), which even tend to be better interpretable than those generated by deep neural networks (Rudin 2019)” (Janiesch et al., 2021)

for time reasons, this thesis focuses (on data preprocessing as well as) on classification approaches only, and does not take into account regression methods.

3.1 Section One

MEG specifics as in Simons study: `https://elifesciences.org/reviewed-preprints/93357` GIFs

3.2 Section Two

Hello

4 Results

5-20 pages

4.1 Section One

Hello

4.2 Section Two

Hello

5 Discussion

5-10 pages
Interpretation of Results
Answering the Forschungsfrage if applicable
bring it back to the literature/theoretical background
implications for further research

Conclusion/Fazit at the end? 3-4 pages
-¿ leave it out at it is the same as above!?!
summary of most important findings
critical refelction

5.1 Section One

Hello

5.2 Section Two

Hello

References

Baillet, S., Mosher, J., & Leahy, R. (2001). Electromagnetic brain mapping. *IEEE Signal Process. Mag.*, *18*, 14-30. doi: 10.1109/79.962275

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., ... Peterson, B. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion*, *10 3*, 377-89. doi: 10.1037/a0018484

Cowen, A. S., & Keltner, D. (2017, 9). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, *114*(38). Retrieved from `https://doi.org/10.1073/pnas.1702247114` doi: 10.1073/pnas.1702247114

Ekman, P., & Cordaro, D. T. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*, 364 - 370. doi: 10.1177/1754073911410740

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013, May). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, *17*(5), 202–209. doi: 10.1016/j.tics.2013.03.006

Gatti, E., Calzolari, E., Maggioni, E., & Obrist, M. (2018). Emotional ratings and skin conductance response to visual, auditory and haptic stimuli. *Scientific Data*, *5*. doi: 10.1038/sdata.2018.120

Gross, J., & Levenson, R. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*, 87-108. doi: 10.1080/02699939508408966

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, *61*, 14 - 5. doi: 10.1177/0008125619864925

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993, 4). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413–497. doi: 10.1103/revmodphys.65.413

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An introduction to statistical learning*. Springer texts in statistics. doi: 10.1007/978-3-031-38747-0

Janiesch, C., Zschech, P., & Heinrich, K. (2021, 4). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695. doi: 10.1007/s12525-021-00475-2

Kensinger, E. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*, 241 - 252. doi: 10.1515/REVNEURO.2004.15.4.241

Kim, H., Lu, X., Costa, M., Kandemir, B., Adams, R. B., Li, J., ... Newman, M. (2018). Development and validation of image stimuli for emotion elicitation (isee): A novel affective pictorial system with test-retest repeatability. *Psychiatry Research*, *261*, 414-420. doi: 10.1016/j.psychres.2017.12.068

Kragel, P., & LaBar, K. (2016). Decoding the nature of emotion in the brain. *Trends in Cognitive Sciences*, *20*, 444-455. doi: 10.1016/j.tics.2016.03.011

Krumnack, U. (2022, June). *The universe of data analytics tools.* Image shown in a practice session for the course 'Machine Learning'. Institute of Cognitive Science, University of Osnabrueck.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-hill New York.

Mitchell, T. M. (2020). *Machine learning - chapter 3 - Generative and Discriminative Classifiers: Naive Bayes and Logistic Regression*. Unpublished book chapter. Retrieved from `http://www.cs.cmu.edu/~tom/mlbook.html`

Mukhamediev, R., Popova, Y., Kuchin, Y., Zaitseva, E., Kalimoldayev, A., Symagulov, A., ... Yelis, M. (2022). Review of artificial intelligence and machine learning technologies: Classification, restrictions, opportunities and challenges. *Mathematics*. doi: 10.3390/math10152552

Polo, E. M., Farabbi, A., Mollura, M., Paglialonga, A., Mainardi, L., & Barbieri, R. (2023). Comparative assessment of physiological responses to emotional elicitation by auditory and visual stimuli. *IEEE Journal of Translational*

Engineering in Health and Medicine, *12*, 171 - 181. doi: 10.1109/JTEHM .2023.3324249

Shaveta, N. (2023, 5). A review on machine learning. *International Journal of Science and Research Archive*, *9*(1), 281–285. doi: 10.30574/ijsra.2023.9.1 .0410

Trepel, M. (2012). *Neuroanatomie: Struktur und Funktion* (5th ed.). Urban & Fischer Verlag/ Elsevier GmbH.

Uhrig, M., Trautmann, N., Baumgärtner, U., Treede, R., Henrich, F., Hiller, W., & Marschall, S. (2016). Emotion elicitation: A comparison of pictures and films. *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.00180

Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience*. doi: 10.1162/jocn.2009.21366

Wardhani, N. W. S., Rochayani, M. Y., Iriany, A., Sulistyono, A., & Lestantyo, P. (2019). Cross-validation metrics for evaluating classification performance on imbalanced data. *2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, 14-18. doi: 10.1109/ IC3INA48034.2019.8949568

Xue, J., Wang, J., Hu, S., Bi, N., & Lv, Z. (2022). Ovpd: Odor-video elicited physiological signal database for emotion recognition. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1-12. doi: 10.1109/TIM.2022 .3149116

Zhou, F., Qu, X., Jiao, R., & Helander, M. (2014). Emotion prediction from physiological signals: A comparison study between visual and auditory elicitors. *Interact. Comput.*, *26*, 285-302. doi: 10.1093/iwc/iwt039

A Appendix

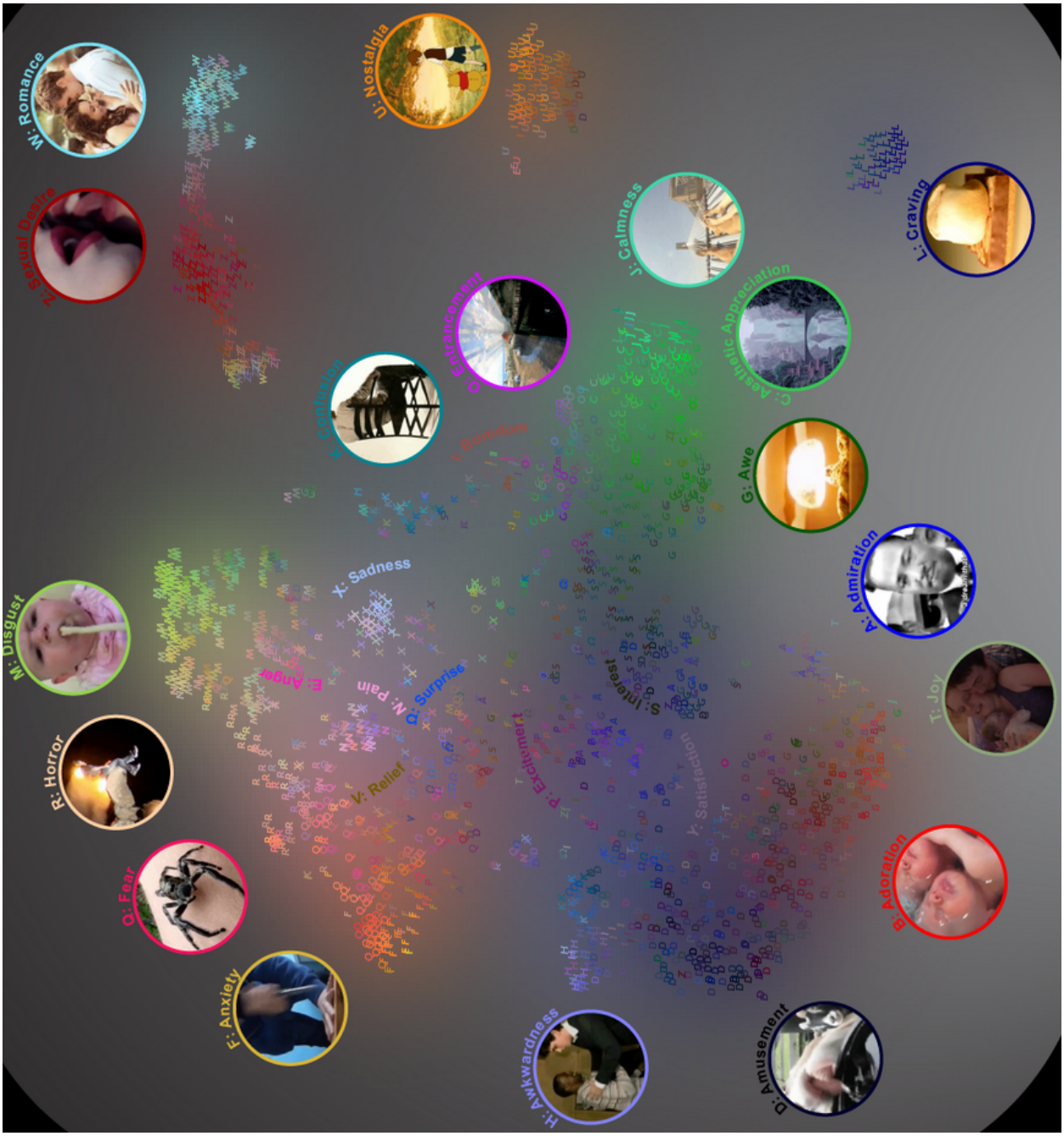

Figure 10: Rotated chromatic map of 27 distinct emotion categories as denoted by Cowen and Keltner (2017). Accessed on 30.08.2024 through `https://s3-us-west-1.amazonaws.com/emogifs/map.html`