# Predicting motoric skills development

Research paper - the development of a prediction model commissioned by Start(v)aardig

**Project**
Start(v)aardig

**Problem owner**
Pim Koolwijk

**Teachers**
Tony Andrioli
Ruud Vermeij
Jeroen Vuurens

**Education**
Minor Applied Data science

**Team**
Motoric

**Authors**

| | |
|---|---|
| Lisa Dumaij | 19049838 |
| Yuliya Litvin | 20161417 |
| Mustafa Mousa | 18068987 |
| Pascal Seegers | 21132844 |
| Joost van Viegen | 18048757 |
| Joep van der zanden | 19134282 |


**Date**
11-01-2022

**Version**

1.0

# 1 Abstract

Motor skills are used in everyday life by walking, running, or bicycling. These skills start to develop at a young age and increase over time especially at the age when they are starting to go to school. One of the biggest problems of today's society is the accessibility of screens such as:  tablets, laptops or other gaming computers at a young age. As a result, children tend to stay at home to take advantage of these technologies rather than playing outside with peers or participating in sports. For this reason, it is important to focus on the motor skill development in children starting at a very young age.

The research described in this paper is about predicting future motor competence in children with historic data. To make sure that children who might need extra help developing their motor skills receive it before it is too late, and the development of their motor skills cannot be sufficiently improved anymore, in which case these problems will affect the adult life or the child.

This research concludes that data science more precisely machine learning can be used to predict if a child is lacking in motor skills a year later. But the perceived motor skills can't be considered as they don't have a great impact on motor skill development.

# 2 Introduction

It has been discovered that nearly half of all children do not exercise enough (SIA, 2019). Children also take the bicycle to school less frequently, stay indoors more often and sit for long hours per day. Because of this, some children's motor skills have developed insufficiently. This development is worrying, because of the physical, emotional, social, and personal value of sport and exercise for children, which is why it is important for children to start being physically active at a young age. That way they will be more likely to experience enjoyment while exercising (Haga, 2009).

Start(v)aardig is a project that started in 2019 and will be conducted until 2023. Within the project, research for the movement skills of children is done, whereas it is investigated how this could be promoted as efficiently and effectively as possible by the neighborhood sports coach. The research is financed by Regieorgaan Praktijkgericht Onderzoek SIA and carried out by a consortium of ten organizations from the sports and exercise sector that is led by The Hague University of Applied Sciences (Alles over Sport, n.d.).

The foundation for these elements is laid by the children from four until six years. It is therefore important to discover motoric deficits at a young age. However, it is not yet clear, which children might have the highest risk to have or develop a motoric deficit, and which features present the highest impact on the motoric skill development. This leads to the research question of this report as follows:

*"How can data science be used to predict whether a child has a chance of developing a lack in motor skills a year later?"*

The main question consists of the following sub-questions:

- *Which biological and socio-demographic variables have an influence on the motoric skills development of children?*
- *Which model has the lowest false negative rate?*
- *Which characteristics have the children with a lack in motor skills in common?*

## 2.1 Related work

Before beginning with this project, research has been conducted to find studies that are somewhat related to this one. Some of those studies dealt with fine, some with gross and some with both fine and gross motor skills development in young children that in most of the studies were between three and six years old. In studies as for example from Wang (2020) and Abdullah et al. (2016), children were tested with various physical exercises to determine their status of motor competence, which demonstrates a similarity to this study that used physical exercises as a testing method as well. Another similarity between existing studies and this study is the investigation of many different features or rather variables that characterize the children, their background, and other related specifics, as well as the importance of each individual feature (Gilbert, 1980b; de Meester et al., 2020b). Further, a differentiation between actual and perceived motor competence was made and explored in the study from de Meester et al. (2020c), which can also be found in this study, as actual and perceived motor competence are viewed separately. Of interest were also studies as from Wang et al. (2020) and Zysset et al. (2018) that included and/ or evaluated parental surveys or rather questionnaires, since this study incorporates this too.

With all the similarities, these existing studies give an interesting insight and knowledge for the topic of motor skills development and a basic understanding in that sphere, which is helpful for this new study, which's goal - of predicting the motor skills development in young children - is still a matter of unknown territory and has never been dealt with in any study before.

# 3 Materials and Methods

## 3.1 Materials

<u>Received data</u>

The final data set consists of several data frames, which are as follows:

- T0 data, this data contents four measurements were taken during the project: competence, motivation, perception, and the BMI (SIA, 2019).
- T1 data, biological data collected during the second measurement moment.
- Questionnaire data, socio-demographic data collected during the first measurement moment (P. Koolwijk).

The final dataset is structured data which contains 1709 rows (children) and 36 columns/features. The goal of this research is to predict whether it is possible to predict future motoric competence using only data available in T0. So, predict if a child lacks motor skills in the future. The data that will be used when training is the proficiency from T0 while the value we are trying to predict is the proficiency from T1.

<u>Data cleaning</u>

Data cleaning is divided in three ways according to Brownlee (2020): basics cleaning, outliers, and imputation.

- *Basics*: Removing redundant columns and rows.
- *Outliers*: Mean and Standard Deviation method.
- *Missing*: KNN, median and mean. Thereby columns that has more than 20% missing values been dropped.

Balance and scale

The balancing and scaling are only done on the training set.

## 3.2 Methods

Feature selection

Features may have little or no correlation with the output variable: MQ-score t1. According to the book van Buijs (2017), a correlation lower than 0.2 is a very weak relationship. These weak links can be filtered out (Schonig et al., 2018). This has been done with the Random Forest model.

Models

Because the Start(V)aardig research is aiming to predict whether someone has motor skills, the best practice is to use a classification model (Minaie, 2021). Because a child is either classified as motor impaired or not.

In a similar study by Gokten and Uyulan (2021), in which the potential for post-traumatic stress syndrome in children is predicted using a classification model. The classification model used is a Random Forest model. The Random Forest Classifier model has also been used in Byeon (2019) research to predict depression and manage the health of caregivers of Alzheimer's patients. As a result, the Random Forest classifier model was the model of choice for this study.

To verify whether the Random Forest classification model was the appropriate model for this study, the results were compared with other classification models, namely: K-nearest neighbors' classifier (Zhang et al., 2018), Decision tree (Burduk & Wozniak, 2012), Gradient Boosting classifier (Hubáček et al., 2018), Bagging classifier (Plaia et al., 2021).

Koehrsen (2019) research shows that hyperparameter tuning must be used to get the best results for each unique prediction model. GridSearchCV is therefore used for this research (RAMADHAN et al., 2017).

Validation

The Hold-out validation will be used by splitting our dataset into a "train", "test" and "validation" set (Novakovic et al., 2017). The dataset is split randomly with a test size of 20%. The train set will be used to learn the model on. After training, the test set will be used to test how well our model will perform on unseen data. Cross-validation is used to estimate the skill of the models (Brownlee, 2020c).

Evaluation

In this project, it is important that the number of false negatives must be narrowed down. It is better to provide help to children that do not actually need the treatment than missing a child that depends on help to improve his state of motor skills, according to Pim Koolwijk (problem owner). For evaluating the model, it is important that the false negative rate is calculated. That can be done with the help of a confusion matrix (Novakovic et al., 2017). An example of this matrix can be seen in figure 3.1.

| | | Predicted class | |
|---|---|---|---|
| | | Negatives | Positives |
| Actual class | Negatives | *a* | *b* |
| | Positives | *c* | *d* |

*Figure 3.1 Confusion Matrix*

# 4 Results

## 4.1 Correlation biological and socio-demographic variables and MQ score

As an overfitting prevention the most valuable features need to be chosen. Figure 4.1 shows the individual importance of each of the features on the target variable. It clearly shows that there are features which don't have an impact on the model performance. Every feature below 0.05 has been removed because otherwise our model will predict with only one feature. Here is room for improvement and taking more distinguishable features into account.

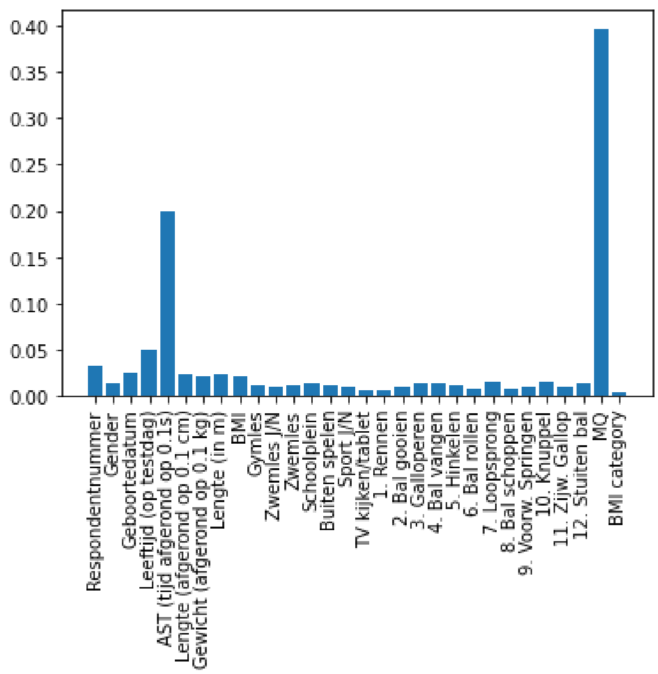


*Figure 4.1 Correlation between features and MQ category*

## 4.2 Models

To achieve the lowest false negative rate, different data preparation methods and models were used.

In the table below the outcome of the different imputation methods is presented. It clearly demonstrates that the mean imputation has the best scores or best positive influence on the models. Therefore, from now on the focus shifted to the mean imputation method, while now focusing only on the results of this imputation method.

*Table 4.1 Imputation methods compared by using the kNN model with binary data*

| Model | Accuracy Train set | Accuracy Test set | False negative rate |
|---|---|---|---|
| Mean imputation | 92,3% | 64,9% | 35,1% |
| Median imputation | 100,0% | 64,9 % | 35,1% |
| KNN imputation | 100,0% | 64,9% | 35,1% |

For researching the best models, a distinction has been made between binary and multilabel classification.

As you can see in the table below the kNN Model is the best model as it does not overfit like the others. Although the Bagging Classifier model had the best accuracy but as it is only 0,3% it can be ignored because this improvement is too low. The interesting thing is that almost all our models scored the same in the false negative rate.

*Table 4.2 Model performances for t0 data with mean outlier removal, mean imputation and binary classification*

| Model | Accuracy Train set | Accuracy Test set | False negative rate |
|---|---|---|---|
| Random forest | 100,0% | 64,9% | 35,1% |
| KNN | 92,3% | 64,9 % | 35,1% |
| Decision Tree | 100,0% | 64,9% | 35,1% |
| Gradient Boost Classifier | 50,0% | 38,8% | 35,1% |
| Bagging Classifier | 98,0% | 65,2% | 34,9% |

The multilabel classification task performed worse than the binary classification task. Comparing this table to table (above) table (below) proves this. These models overfit.

*Table 4.3 Accuracy scores of models using multilabel classification*

| Model | Accuracy Train set | Accuracy Test set |
|---|---|---|
| Random forest | 100,0% | 8,6% |
| KNN | 62,5% | 0,3% |
| Decision Tree | 100,0% | 20,2% |

For model evaluation and to look for improvements a cross-validation is used. The scoring method for the cross-validation is the accuracy. As the kNN model performed best the cross-validation is done on the kNN model.

*Table 4.4 The 10-Fold cross-validation using the accuracy scoring method*

| N-Fold | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Accuracy | 90,0% | 85,6% | 93,1% | 89,9% | 88,1% | 91,8% | 92,5% | 91,2% | 90,6% | 88,1% |

*Table 4.5 The mean and standard deviation for the 10-Fold cross-validation*

| KNN binary model | Accuracy |
|---|---|
| Mean | 90,1% |
| Standard deviation | 2,2% |

## 4.3 Characteristics of a kid with a lack in motor skills

As you can see in the figure 4.1 the given features in the dataset don't have that much of an impact on the motor skill. Therefore, there isn't a pattern or characteristics of a child with a lack of motor skills.

# 5 Discussion

The results of this study can only be used for the StartVaardig project or in a study in which the same tests are performed in the manner as stated in the SIA report from 2019.

In the study, the times can differ between T1 and T0, these time variants will be included as a feature in the study. A time of six months (from September to the end of January) has been set aside for this research.

Another discussion is that not all participants are between the age of 4 and 6, this means that the results maybe not apply to them. Thereby is a large part of the data being incomplete. As a result, not all data could be used, resulting in too little data. This can negatively affect the operation of the algorithms.

# 6 Conclusion & recommendations

To answer our main question, we must first answer our sub-questions.

For our first sub-question it could be concluded that not all the data we received was usable. We discovered for example that for our model the perceived motor competence wasn't as helpful as we had thought at first. The data from the questionnaire from T0 data wasn't complete enough to be usable. This didn't leave a lot of data to train on, which might explain why our models are overfitting.

While researching we stumbled across data from the Centraal Bureau voor de Statistiek but we couldn't merge it to our t0 data as the CBS data was too complex. For future work it might be helpful to investigate data from the Centraal Bureau voor de Statistiek.

Our results show clearly that using a binary classification works best for our study because we are only trying to predict if a child will lack in motor competence, the different categories do not matter as much.

Although research suggests using a Random Forest model (Gokten and Uyulan, 2021) after running and evaluating different models we concluded that the Random Forest isn't appropriate for our research and therefore must be dropped. In table (binary classification) it is pictured that the k-nearest-neighbors model performed best for our research. The Bagging Classifier has the lowest false negative rate however this model overfits worse than kNN (as pictured in table binary classification) so we decided to pursue the kNN model to prevent this.

One possible reason for overfitting might also be that we used t0 data for predicting the MQ category of t1 because the learning curve for motor skill will get steeper for children with good motor skills at some point while it will flatten for children with bad motor skills (Haga, 2009).

In order to be able to correctly predict the future motoric competence of children, there might need to be more variance in the data of the children with low motoric skills. A year might also not be long enough to get a good trend of the score per child. A similar study showed that there is a significant difference after 32 months (Haga, 2009).

We found in our results that there are no common characteristics in our dataset. This might be because perceived motor competence does not have an impact on motor skills. Also, because children below the age of eight do not have a good self-perception of their skills (Morano, 2020).

## References


Alles over Sport. (n.d.). *Start (V)aardig*. Allesoversport.nl. Retrieved October 25, 2021, from https://www.allesoversport.nl/startvaardig/

Annette Brons, Antoine de Schipper, Svetlana Mironcika, Huub Toussaint, Ben Schouten, Sander Bakkes, Ben Kröse (2021, April). *Assessing Children's Fine Motor Skills With Sensor-Augmnted Toys: Machine Learning Approach.* JMIR Publications. Retrieved December 16, 2021, from https://www.jmir.org/2021/4/e24237

Annina E. Zysset, Tanja H. Kakebeeke, Nadine Messerli-Bürgy, Andrea H. Meyer, Kerstin Stülb, Claudia S. Leeger-Aschmann, Einat A. Schmutz, Amar Arhab, Valentina Ferrazzini, Susi Kriemler, Simone Munsch, Jardena J. Puder, Oskar G. Jenni (2018, February). *The validity of parental reports on motor skills performance level in preschool children: a comparison with a standardized motor test.* NCBI. Retrieved December 16, 2021, fromhttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5899107/

Borhannudin Abdullah, Wan Azira Abd Aziz, Aminuddin Yusof (2016, October). *Level of motor skill development of preschool students.* Journal of Physical Education and Sport (JPES). Retrieved December 17, 2021, from https://efsupit.ro/images/stories/3%20September2016/art%20175.pdf

Brownlee, J. (2020). *Data Preparation of Machine Learning*. Jason Brownlee.

Brownlee, J. (2020b, August 20). *How to Choose a Feature Selection Method For Machine Learning*. Machine Learning Mastery. Retrieved November 11, 2021, from https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/

Brownlee, J. (2020c, August 2). *A Gentle Introduction to k-fold Cross-Validation*. Machine Learning Mastery. https://machinelearningmastery.com/k-fold-cross-validation/

Buijs, A. (2017). *Statistiek om mee te werken* (10de editie). Noordhoff.

Burduk, R., & Wozniak, M. (2012). *Different decision tree induction strategies for a medical decision problem. Open Medicine*, 7(2), 183–193. Retrieved November 11, 2021, from https://doi.org/10.2478/s11536-011-0142-x

Byeon, H. (2019). *Developing a random forest classifier for predicting the depression and managing the health of caregivers supporting patients with Alzheimer's Disease*. Technology and Health Care, 27(5), 531–544. Retrieved November 12, 2021, from https://doi.org/10.3233/thc-191738

Chairilsyah, D. (2019). *Web-Based Application to Measure Motoric Development of Early Childhood*. JPUD - Jurnal Pendidikan Usia Dini, 13(1), 1–14. Retrieved Novermeber 12, 2021, from https://doi.org/10.21009/10.21009/jpud.131.01

Centraal Bureau voor de Statistiek. (2019, December 10). *Inkomensverdeling per postcodegebied (PC4), 2017*. Retrieved December 17, 2021, from https://www.cbs.nl/nl-nl/maatwerk/2019/50/inkomensverdeling-per-postcodegebied--pc4---2017

Centraal Bureau voor de Statistiek. (2021a, September 17). *Bevolking; geslacht, migratieachtergrond, viercijferige postcode, 1 januari.* Retrieved December 17, 2021, from https://www.cbs.nl/nl-nl/cijfers/detail/83503NED

Centraal Bureau voor de Statistiek. (2021b, December 17). *Kerncijfers per postcode.* Retrieved December 17, 2021, from https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

De Meester, A., Barnett, L.M., Brian, A. *et al.* The Relationship Between Actual and Perceived Motor Competence in Children, Adolescents and Young Adults: A Systematic Review and Meta-analysis. *Sports Med* **50,** 2001–2049 (2020). . Retrieved December 15, 2021, from https://doi.org/10.1007/s40279-020-01336-2

Dmitrievsky, M. (2018, July 6). *The abstract description of the Random Forest algorithm*. MQL5. Retrieved October 17, 2021, from https://www.mql5.com/en/articles/3856

Dmitrievsky, M. (2018, July 6). *RANDOM DECISION FOREST IN REINFORCEMENT LEARNING*. mql5. Retrieved December 15, 2021, from https://www.mql5.com/en/articles/3856

Gokten, E. S., & Uyulan, C. (2021). *Prediction of the development of depression and post-traumatic stress disorder in sexually abused children using a random forest classifier. Journal of Affective Disorders, 279, 256–265*. Retrieved December 17, 2021, from https://doi.org/10.1016/j.jad.2020.10.006

Gilbert, J. (1980). *An Assessment of Motor Music Skill Development in Young Children*. Journal of Research in Music Education, 28(3), 167–175. Retrieved December 17, 2021, from https://doi.org/10.2307/3345234

Hubáček, O., ŠOurek, G., & ŽElezný, F. (2018). *Learning to predict soccer results from relational data with gradient boosted trees.* Machine Learning, 108(1), 29–47. Retrieved December 15, 2021, from https://doi.org/10.1007/s10994-018-5704-6

Kazil, J., & Jarmul, K. (2016). *Data Wrangling with Python: Tips and Tools to Make Your Life Easier* (1st ed.). O'Reilly Media.

Koehrsen, W. (2019, December 10). *Hyperparameter Tuning the Random Forest in Python - Towards Data Science*. Medium. Retrieved December 21, 2021, from https://towardsdatascience.com/hyperparameter-tuning-the-random-forest-in-python-using-scikit-learn-28d2aa77dd74

Kuhn, M., & Johnson, K. (2019). *Feature Engineering and Selection: A Practical Approach for Predictive Models (Chapman & Hall/CRC Data Science Series)* (1st ed.). Chapman and Hall/CRC.

Martin-Ruiz, M. L. (2015). *Foundations of a Smart Toy Development for the Early Detection of Motoric Impairments at Childhood.* International Journal of Pediatric Research, 1(2). Retrieved October 11, 2021, from https://doi.org/10.23937/2469-5769/1510011

Meester, D. A. (2020, September 24). *The Relationship Between Actual and Perceived Motor Competence in Children, Adolescents and Young Adults: A Systematic Review and Meta-analysis.* SpringerLink. Retrieved December 8, 2021, from https://link.springer.com/article/10.1007/s40279-020-01336-2?error=cookies_not_supported&code=37b734cf-1842-49b4-a1db-a4832a112243

Minaie, N., PhD. (2021, December 10). *The Data Scientist's Guide to Selecting Machine Learning Predictive Models in Python*. Medium. Retrieved December 15, 2021, from https://towardsdatascience.com/the-beginners-guide-to-selecting-machine-learning-predictive-models-in-python-f2eb594e4ddc

Monika Haga, Physical Fitness in Children With High Motor Competence Is Different From That in Children With Low Motor Competence, *Physical Therapy*, Volume 89, Issue 10, 1 October 2009, Pages 1089–1097, https://doi.org/10.2522/ptj.20090052

Morano M, Bortoli L, Ruiz MC, Campanozzi A, Robazza C (2020) Actual and perceived motor competence: Are children accurate in their perceptions? PLoS ONE 15(5): e0233190. https://doi.org/10.1371/journal.pone.0233190

Novakovic, J. D. J., Veljovic, A., Ilic, S. S., Papic, Z., & Tomovic, M. (2017). *Evaluation of Classification Models in Machine Learning*. UAV. Retrieved December 1, 2021, from https://uav.ro/applications/se/journal/index.php/TAMCS/article/view/158/126

Plaia, A., Buscemi, S., Fürnkranz, J., & Mencía, E. L. (2021). *Comparing Boosting and Bagging for Decision Trees of Rankings.* Journal of Classification. Retrieved December 2, 2021, from https://doi.org/10.1007/s00357-021-09397-2

RAMADHAN, M. M., SITANGGANG, I. S., NASUTION, F. R., & GHIFARI, A. (2017). *Parameter Tuning in Random Forest Based on Grid Search Method for Gender Classification Based on Voice Frequency.* DEStech Transactions on Computer Science and Engineering, cece. Retrieved December 17, 2021, from https://doi.org/10.12783/dtcse/cece2017/14611

Sander J., Schipper A., Brons A., Mironcika S., Toussaint H., Schouten B., Kröse B. (unknown). *Detecting delays in motor skill development of children through data analysis of a smart play device.* Unknown. Retrieved October 11, 2021, from https://digitallifecentre.nl/redactie/resources/finalpaperfinal.pdf

Sia. (2019). *Aanvraagformulier RAAK-PRO Start (V)aardig -2018*. Nationaal Regieorgaan Praktijkgericht Onderzoek SIA.

Schonig, S., Jasinski, R., Ackermann, L., & Jablonski, S. (2018, January). *Deep Learning Process Prediction with Discrete and Continuous DataFeatures*. ResearchGate. Retrieved December 17, 2021, from https://doi.org/10.5220/0006772003140319

Wang, H., Chen, Y., Liu, J., Sun, H., & Gao, W. (2020). *A Follow-Up Study of Motor Skill Development and Its Determinants in Preschool Children from Middle-Income Family*. BioMed Research International, 2020, 1–13. Retrieved December 1, 2021, from https://doi.org/10.1155/2020/6639341

Zhang, S., Li, X., Zong, M., Zhu, X., & Wang, R. (2018). *Efficient kNN Classification with Different Numbers of Nearest Neighbors.* IEEE Transactions on Neural Networks and Learning Systems, 29(5), 1774–1785. Retrieved December 2, 2021, from https://doi.org/10.1109/tnnls.2017.2673241