Variable selection in gender and age decision-making for traumatic spine and thoracic pathologies after various accidents: MARS

Kerem Gencer
Afyonkarahisar Health Sciences University
Gülcan Gencer (gencergulcan@gmail.com)
Afyonkarahisar Health Sciences University
Ahmet Dumanlı
Afyonkarahisar Health Sciences University

Article

Keywords: Thorax, vertebra, MARS, variable selection, traffic collision, autopsy

Posted Date: January 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2404787/v1

Abstract

Trauma is a condition that affects the body's structure and results from outside factors. After heart disease and cancer, it is the most common cause of death across all age categories. For a variety of causes, people are routinely exposed to traumatic vertebral, thoracic pathologies and rib fractures. Ribs can be harmed by simple falls, impacts, and blunt injuries as well as broken due to car accidents and falling from a height. Magnetic resonance imaging or computed tomography are used to diagnose these fractures. In this study, non-linear complex methods were used to categorize gender and age by utilizing thoracic pathologies, fractures or cracks in the body as a result of traffic accidents or falling from a height, which have the feature of being a case in forensic issues. Variables were selected to identify the most important data. MARS (Multivariate Adaptive Regression Spline) method of variable selection. Although autopsy should be utilized in these situations, complex regression methods is intended to have an impact on quick and accurate decision-making about events in order to speed up or direct the process in the field of forensic medicine. As a result, the effectiveness of the experts' subsequent predictions will be increased by the preliminary findings produced by real-world data and artificial intelligence algorithms or complex non-linear regression problems.

1. Introduction

For children and young adults between the ages of 5 and 29, traffic-related injuries are the main cause of death. An estimated 1.3 million people each year pass away in traffic-related incidents. Men are more likely than women to get in car accidents starting at a young age. Young men under the age of 25 account for nearly three times as many road traffic fatalities (73%) as young women. [1].

In Turkey in 2021, there were 5 000 362 fatal road accidents, but there were also 274 000 615 injuries. 2,421 individuals died at the scene of the 187 thousand 963 death and injury traffic accidents, and 2,941 people were hurt and sent to medical facilities within 30 days as a result of the accident's cause and effect. A residential location accounted for 49.3% of fatalities and 72.0% of injuries, whereas a non-residential region accounted for 50.7% of deaths and 28.0% of injuries [2]. Following a car collision, trauma is a condition that disturbs the body's structural order as a result of external forces [3]. 40% of all trauma cases involve thoracic trauma, the third most frequent type of injury overall [3, 4, 5]. Thoracic injuries make up one-third of all trauma emergency and hospital admissions [5, 7]. A tomographic image of the vertebral and rib fractures is shown in Fig. 1.

Ribs can be harmed by simple falls, impacts, and blunt injuries as well as broken due to car accidents and falling from a height. The crime scene is investigated as soon as the accident occurs on the road. Physicians are supposed to take part in gathering any biological evidence that may exist, assess if the incidence was truly an accident, and decide whether an autopsy is necessary if a deceased person is present [8]. The Turkish Penal Code states in Article 280 that "if physicians, dentists, pharmacists, midwives, nurses and other healthcare providers fail to report the situation to the competent authorities or delay in doing so, despite discovering a sign that a crime has been committed while performing their

duties, shall be punished with imprisonment of up to one year." Because of this, individuals who visit the emergency department or health facility; Situations including traffic accidents, falls, assaults, and work accidents should be treated as legal cases and reported in writing or verbally to the appropriate authorities [9].

However, in terms of both criminal and civil law, forensic medicine's determination of age and gender is significant. The least amount of mistake should be used when defining anatomical features [10]. Identification will continue to be one of the most important aspects in forensic cases, according to studies done so far.

A thorough investigation on the density and burden of fractures was also carried out for the first time in 2019 because bone fractures are a hazard for world health. According to age and sex, Fig. 2 shows the global fracture incidence rate for each anatomical location from 1990 to 2019 [11].

In many traffic collisions, especially those involving large cars, there are unrecognizably dead bodies [12, 13]. To date, fractures have been found using techniques like MR and CT for a variety of causes. However, in this study, it has been proposed to estimate gender and age by looking for fractures or cracks in the body by combining deep learning and machine learning techniques. First, variable selection was done, and then 150 features were chosen for both gender and age requirements. Making the data noiseless brought the feature reduction step before classification to a successful conclusion. For feature selection, a data collection with 251 samples overall was employed.

The study's variables, feature selection are all detailed in Chap. 2 of this article. Evaluation metrics, feature reduction, and overall analysis are the main topics of Chap. 3. Section 4 of the report discusses the findings. Chapter 5 presents the findings.

2. Material And Methods

Thorax computed tomography scans were examined to determine which patients had spinal, rib fractures and thoracic pathologies. In forensic medicine, cases such as traffic accidents, falls, assaults, work accidents are accepted as forensic cases. Therefore, the study included patients with rib fractures or thoracic vertebral fractures after a car collision and fall from a height. On a chest computed tomography, the degree of rib fractures, displacement, or non-displacement, were identified. Patients with thoracic vertebral fractures had their fracture levels, kinds (corpus fracture, transverse, and spinous process fractures) and thoracic pathologies identified. For statistical calculations, the R 4.0.2 software environment was employed. Patients provided both written and verbal informed consents for the study, which were prepared in accordance with the Helsinki Declaration and approved by the clinical research ethics committee of the Afyonkarahisar Health Sciences University (Ethics Committee Approval Date - No: 02.12.2022-2022/16). The study was single-centered in design.

2.1. Featuare Selection

A technique for identifying and removing unimportant participants from a data collection is feature selection. As a result of this procedure, the study moves forward using the information that best explains the topic under investigation. As a result, multidimensionality's complexity is eliminated. The literature describes three feature selection techniques: filtering, wrapping, and embedding. One of the wrapping approaches, MARS, was employed in this work to pick features [14].

2.1.1. MARS

MARS is one of the new non-parametric regression methods. It was created in the early 1990s by Jerome H. Friedman [15], a statistician at Stanford University. MARS is a hybrid of Recursive Partitioning Regression (RPR) and Projection Pursuit Regression (PPR) (Banks 2001). This regression procedure is designed to accommodate both binary and continuous output variables. MARS is a versatile, precise, and quick technique. In contrast to linear methods, it considers subsets of variables [16]. This method, which is used for classification with a categorical output variable, has a wide range of applications because it is very flexible, accurate, and fast [17].

For $(x \in \mathcal{R})$, the following is the basis function:

$$(x-t)_+ = \begin{cases} x-t, if x > t \\ 0, otherwise \end{cases} \text{ ve } (t-x)_+ = \begin{cases} t-x, if x < t \\ 0, otherwise \end{cases} \quad (1)$$

In the expression above, t represents the point at which every function becomes piecewise linear at the value t. Figure 4 depicts a spline function's inflection point. The goal of this inflection is to obtain the projected pair x_j via nodes with the value x_{ij}. The sum of all basis functions is depicted below[18]. Figure 3 shows a graph of the piecewise cubic type of the MARS model.

$$C = \left\{(x_J - t)_+, (t - x_J)_+\right\}$$

2

The maximum model obtained initially is cropped with the second step, the backtracking algorithm, and the best model is discovered by eliminating the least effective variables one by one in the MARS model construction process. The submodel with the best approximation is determined by comparing the submodels obtained in this process using the Generalized Cross-Validation (GCV) criterion. The GCV method is described below.

$$GCV = \frac{\sum_{i=1}^{N} \left(y_i - \widehat{f}_{\lambda}(x_i)\right)^2}{\left(1 - M(\lambda)/N\right)^2}$$

3

$M(\lambda)$ and N denote the number of effective parameters and the number of observations, respectively. $M(\lambda)$ is found in this expression as $M(\lambda) = r + cK$ where r and K denote the number of

independent basis functions and the number of nodes chosen in the incremental part, respectively.

3. Case Study

251 patients who experienced various trauma cases traffic accidents and fall injuries were included in the study. In the model constituting the data set, 63.35% occurred as a result of falling from a height, 36.65% as a result of traffic accidents. These patients had thoracic vertebral fractures as well as rib fractures or rupture, injury, bleeding, and other conditions in the lung. According to fracture or fracture status, it was shown that the degrees and types of fractures (corpus fracture, transverse, and spinous process fractures) in patients with spinal and/or rib fractures and in patients with thoracic vertebral fractures varied. 150 different variables were taken into account. Varaible status was coded with a 1 rather than a 0.

3.2. Feature reduction

For two variables in this investigation, feature reduction was done. Age and gender are these factors. A total of 150 features were reduced to 9 features using the MARS feature selection method for the age variable. These are Lung Parenchymal laceration, Lung left lower lobe lobe atelectasis, right 7th rib displaced fracture, right 10th rib displaced fracture, left 6th rib displaced difference, left 8. rib non-displaced fracture, left 9th rib non-displaced fracture, C7 vertebra transverse process fracture, T10 vertebra transverse fracture.

A total of 150 features were reduced to 5 features using the MARS feature selection method for the gender variable. These were identified as right 3rd rib displaced fracture, right 7th rib displaced fracture, left 5th rib non-displaced fracture, left 9th rib non-displaced fracture, T2 vertebra corpus fracture

3.3 Analysis Results

Table 1 shows demographic data. Male make up 148% of the respondents, and female make up 103 percent. In the data collection, there are 40 people over the age of 71, 114 people between the ages of 45–70 and 97 people between the ages of 16–44.

Table 1
Features of the population

Variables		Frequency	Mean
Sex	Male	148	45,5
	Female	103	55,2
Age	16–44	97	
	45–70	114	49,25
	71 and above	40	
Total		**251**	

Table 2 shows the treatment method applied according to the etiopathology of the diseases. As a result of traffic accidents, surgical methods were used in 58.7% and medical methods were used in 41.3%. As a result of falling from a height, 71.1% of the patients were treated with surgical methods, while 28.9% were treated with medical methods.

Table 2
Treatment method according to the disease etiopathology

Treatment method		Traffic Accidents	Falls from height	Total
Surgical	Count	54	113	169
	% within etiopathology	58.7%	71.1%	66.8%
Medical	Count	38	46	84
	% within etiopathology	41.3%	28.9%	33.2%
	Count	92	159	251

In the MARS feature selection, 150 features were reduced to 5 features for sex and 9 features for age, as shown in Table 3. Lung parenchymal laceration, Lung left lower lobe atelectasis, right 7th rib displaced fracture, right 10th rib displaced fracture, left 6th rib displaced difference, left 8th rib non-displaced fracture, left 9th rib non-displaced fracture, C7 vertebra transverse process fracture, T10 vertebra transverse fracture for age.

Table 3
Selected Features Performances for gender and age

Feature Selection Techniques	Number of features Selected	Selected Features (MARS)
Sex	5	Right 3rd rib displaced fracture, right 7th rib displaced fracture, left 5th rib non-displaced fracture, left 9th rib non-displaced fracture, T2 vertebra corpus fracture
Age	9	Lung Parenchymal laceration, Lung left lower lobe lobe atelectasis, right 7th rib displaced fracture, right 10th rib displaced fracture, left 6th rib displaced difference, left 8. rib non-displaced fracture, left 9th rib non-displaced fracture, C7 vertebra transverse process fracture, T10 vertebra transverse fracture.

The features selected by the algorithms are given in Fig. 4. Commonly selected are left 9th rib non displaced and right 7th rib displaced fracture.

It is suggested in this study to increase the use of complex non-linear regression methods in the field of health management and services. In our study, we found that the doctors' decisions are consistent with the outcomes of MARS. This demonstrates that complex non-linear regression can be used to reduce human-induced errors in diagnosis and treatment processes, as well as to aid in medical decision-making.

4. DISCUSSION

It is very important to determine the age and gender of corpses that are not recognized in the discipline of forensic medicine. In this study, while determining the first criteria to be considered when estimating gender and age, criteria such as fractures or cracks in the body, bruises and injuries were reviewed. Those who experienced trauma after a traffic accident and fell from a height were included in the study. These patients have possible thoracic vertebral fractures and thoracic pathologies as well as rib fractures. The factors seen in these patients were first reduced, and then the most important variables were found. Cases after traffic accidents and falls from height, which can be the subject of forensic medicine, formed the basis of the data used in the study. In this study, features that should be considered in age and gender estimation were estimated using a feature technique. It has been suggested that deep learning, artificial intelligence and machine learning algorithms or complex nonlinear regression problems in forensic science can be used in forensic applications.

5. Conclusions

A person may need recognition and separation from others for any reason. Both the government and the person's relatives attach great importance to this scenario. In these cases, forensic medicine institutions

provide assistance. A person's identity consists of all the characteristics that allow identification and distinguish him from other people. A person may need recognition and separation from others for any reason. Using the complex nonlinear regression method on photographs of patients with spine and/or rib fractures and patients with thoracic vertebral fractures, this study identified the first features to consider when determining the age and sex of patients. X-rays were taken first. Patients with spine and/or rib fractures and patients with thoracic vertebral fractures were classified in a database by looking at these photographs. As a result, analysis of fractures or cracks seems to offer valuable information as a starting point for research. However, it is clear that the purpose of such a classification is to arrive at a general conclusion rather than a definitive answer. Accuracy rates are expected to serve as a guideline for the study to have appropriate evidential value. This research seeks to guide future research. As a result, machine learning algorithms or complex nonlinear regression problems used in forensics could help open the door to more experimental work in other disciplines. It is estimated that these people will be a new guide to the authorities when it is determined that the people who suddenly disappeared are buried in a place years later.

Declarations

Acknowledgements

We thanks all participants of the Afyonkarahisar Health Sciences University Hospital Chest Surgery Department

Authors' contributions

KG and GG was involved in the study's design and execution, as well as data analysis, manuscript drafting, and critical discussion. GG and KG helped with the study's design and execution, data analysis, manuscript drafting, and critical discussion. AD was involved in the study design and execution, and KG, GG, and AD were involved in the data analysis and manuscript drafting. AD all helped with laboratory testing and manuscript writing. The final manuscript was read and approved by all authors.

Ethics approval and consent to participate

All participants signed written informed consent forms and the study. Approval was obtained from Afyonkarahisar Health Sciences University Clinical Research Ethics Committee (Approval date – No: 02.12.2022–2022/16).

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Competing interests

The author declares no conflict of interest related to this article.

Funding sources

The authors declares that this study has received no financial support.

References

1. Organization, W.H. *Road traffic injuries*. 2022; Available from: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.
2. Kurumu, T.İ., *Karayolu Trafik Kaza İstatistikleri, 2021*. 18 Mayıs 2022.
3. Haberal, M., et al., *Traumatic hemothorax: Analysis of 108 cases*. Journal of Experimental and Clinical Medicine, 2013. **30**(1): p. 31–33.
4. Apilioğulları, B., et al., *Retrospektive analysis of 48 cases with thoracic trauma*. Anatol J Med Sci, 2015. **1**(1): p. 14–18.
5. Ceran, S., et al., *Chest trauma in children*. European journal of cardio-thoracic surgery, 2002. **21**(1): p. 57–59.
6. Büyükkarabacak, Y., et al., *Toraks travmalı hastalarda yandaş travmalar: mortalite ve morbidite üzerine etkileri*. Black Sea Journal of Health Science, 2019. **2**(3): p. 78–84.
7. Tülüce, K. and G. Altuntaş, *Travmatik Pnömotorakslı 127 Hastanın Değerlendirilmesi: Tek Merkez Deneyimi*. Sakarya Tıp Dergisi, 2020. **10**(4): p. 655–660.
8. Çetin, G. and A. Özaslan, *Trafik kazasına bağlı yaralanmalar*, Z. Soysal and C. Çakalır, Editors. 1999: Adli Tıp İstanbul: İstanbul Üniversitesi
9. TCK, *5237 Sayılı Türk Ceza Kanunu (TCK), 280. Madde*.
10. Aktaş, E., *kostaların Sternal Uç Kemik Morfolojisinde Yaşa İlişkin Progressif Değişikliklerin Kişinin Öldüğü Zamanki Yaşının Saptanmasında Kullanılabilirliği.* Specialization Thesis, Faculty of Medicine, İzmir, 1997.
11. Collaborators, G.F., *Global, regional, and national burden of bone fractures in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019*. Lancet Healthy Longev, 2021. **2**(9): p. e580-e592.
12. İHA. *Dumanlar gökyüzünü kapladı! Çin'de petrol tankeri kamyonla çarpıştı*. 2022; Available from: https://www.cnnturk.com/dunya/dumanlar-gokyuzunu-kapladi-cinde-petrol-tankeri-kamyonla-carpisti.
13. Cİhan. *Otoyolda 'U dönüşü" bir aileyi yok etti*. 2010; Available from: https://www.haber7.com/guncel/haber/632235-otoyolda-u-donusu-bir-aileyi-yok-

14. Chandrashekar, G. and F. Sahin, *A survey on feature selection methods*. Computers & Electrical Engineering, 2014. **40**(1): p. 16–28.
15. Friedman, J. H., 1991, Multivariate adaptive regression splines (with discussion), The Annals of Statistics, 19(1), 1–141.
16. Muñoz, J. and Felicísimo, Á., 2004, Comparison of statistical methods commonly used in predictive modelling, Journal of Vegetation Science, 15, 285–292.
17. Mukhopadhyay, A. and Iqbal, A., 2009, Prediction of mechanical property of steel strips using multivariate adaptive regression splines, Journal of Applied Statistics 36(1), 1–9.
18. Put, R., Xu, Q. S., Massart, D. L. and Vander Heyden, Y., 2004, Multivariate adaptive regression splines (MARS) in chromatographic quantitative structure–retention relationship studies, Journal of Chromatography A, 1055(1), 11–19.
19. Kartal Koc, E., & Bozdogan, H. (2015). Model selection in multivariate adaptive regression splines (MARS) using information complexity as the fitness function. Machine Learning, *101*(1), 35–58.

Figures

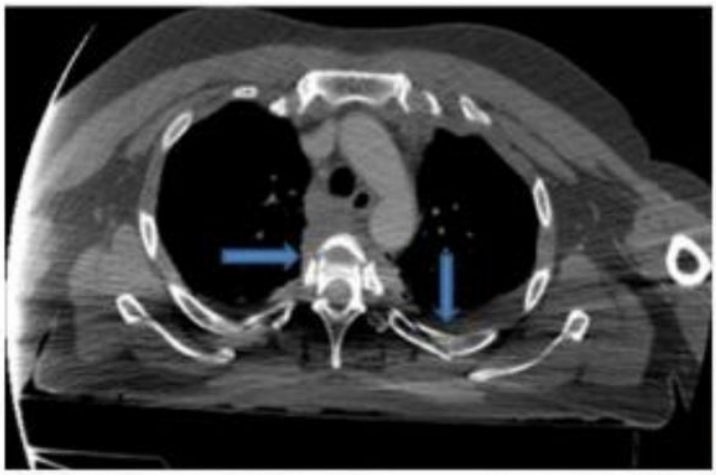

Figure 1A. Fracture of the 5th vertebra corpus and fracture of the left 5th rib

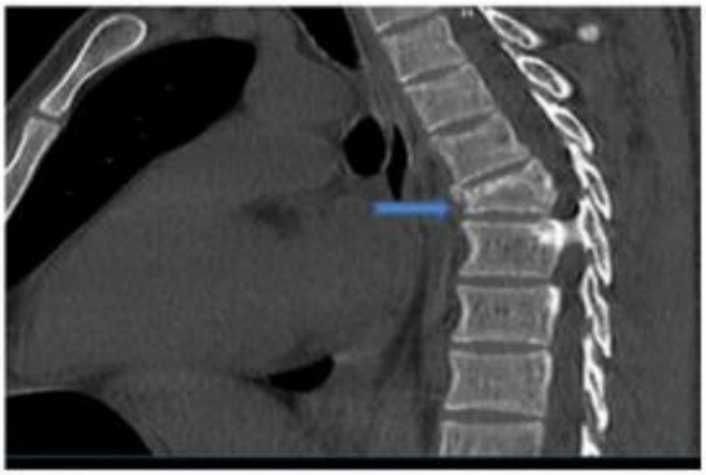

Figure 1B. T6 vertebra compression fracture

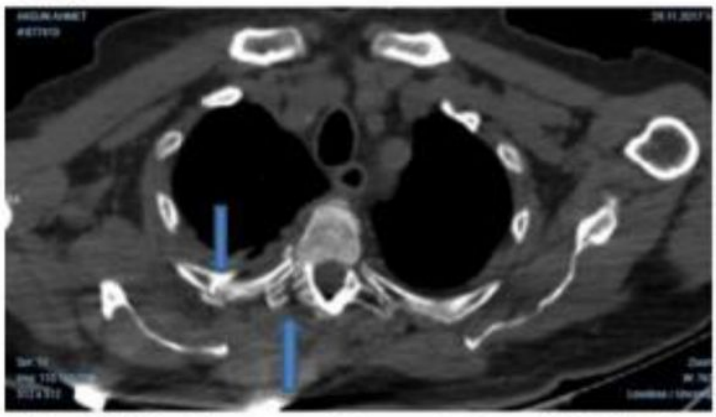

Figure 1C. 4th vertebra transversal process fracture and right 4th rib deplace fracture

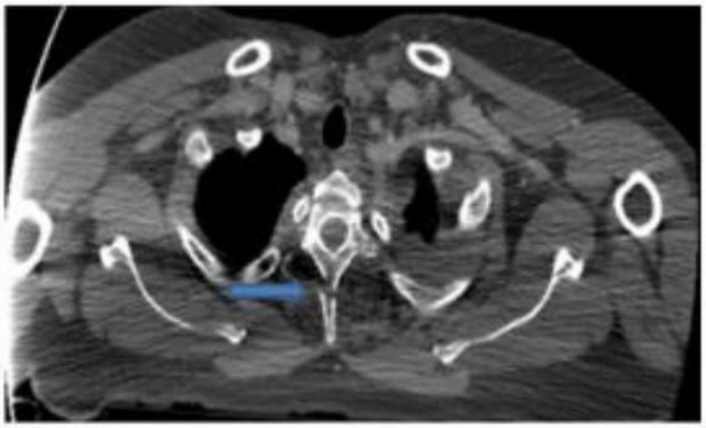

Figure 1D. T2 vertebra spinous process fracture

Figure 1

Tomography pictures of rib and vertebral fractures

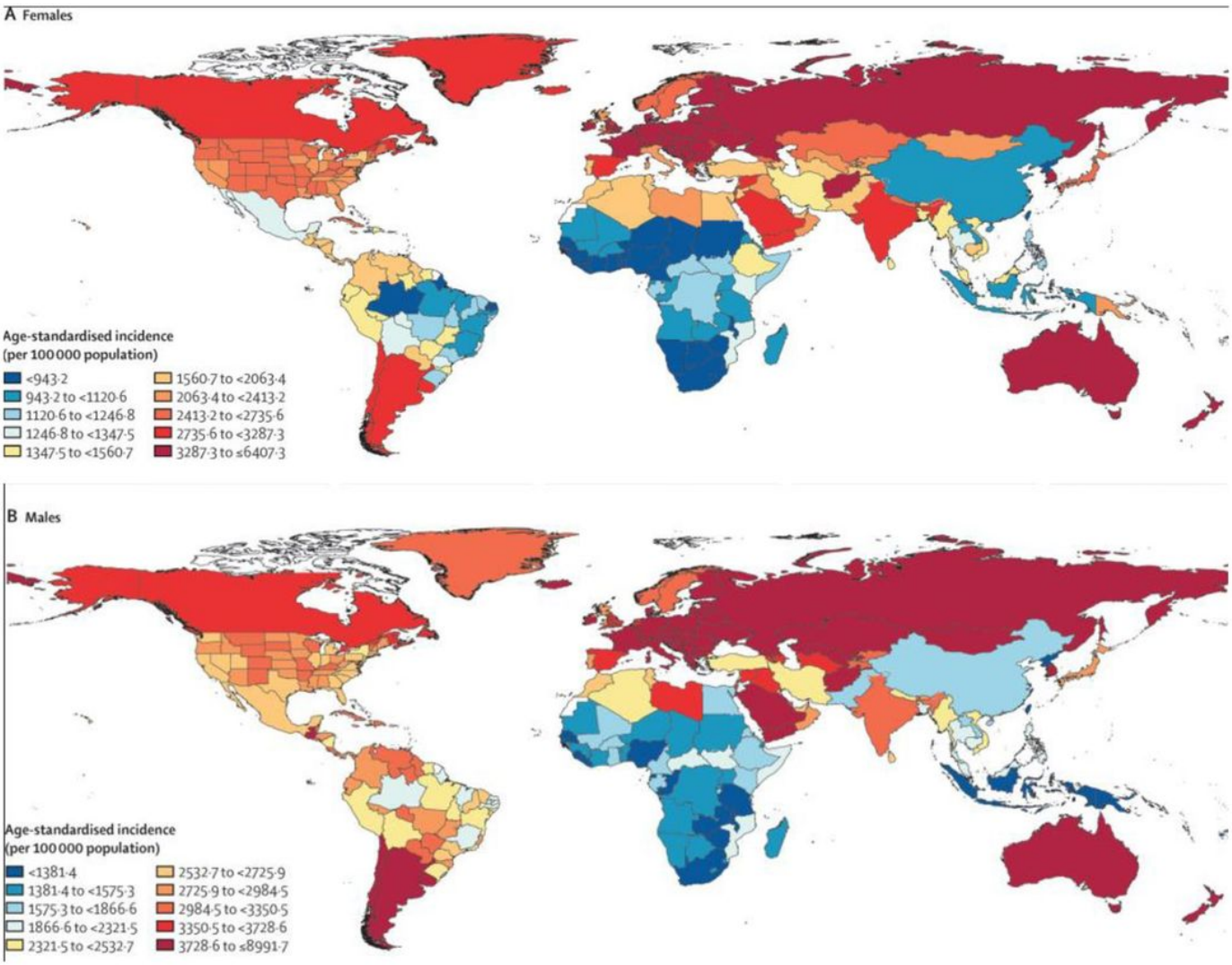

Figure 2

Age-standardised incidence of fractures for females (A) and males (B), 2019

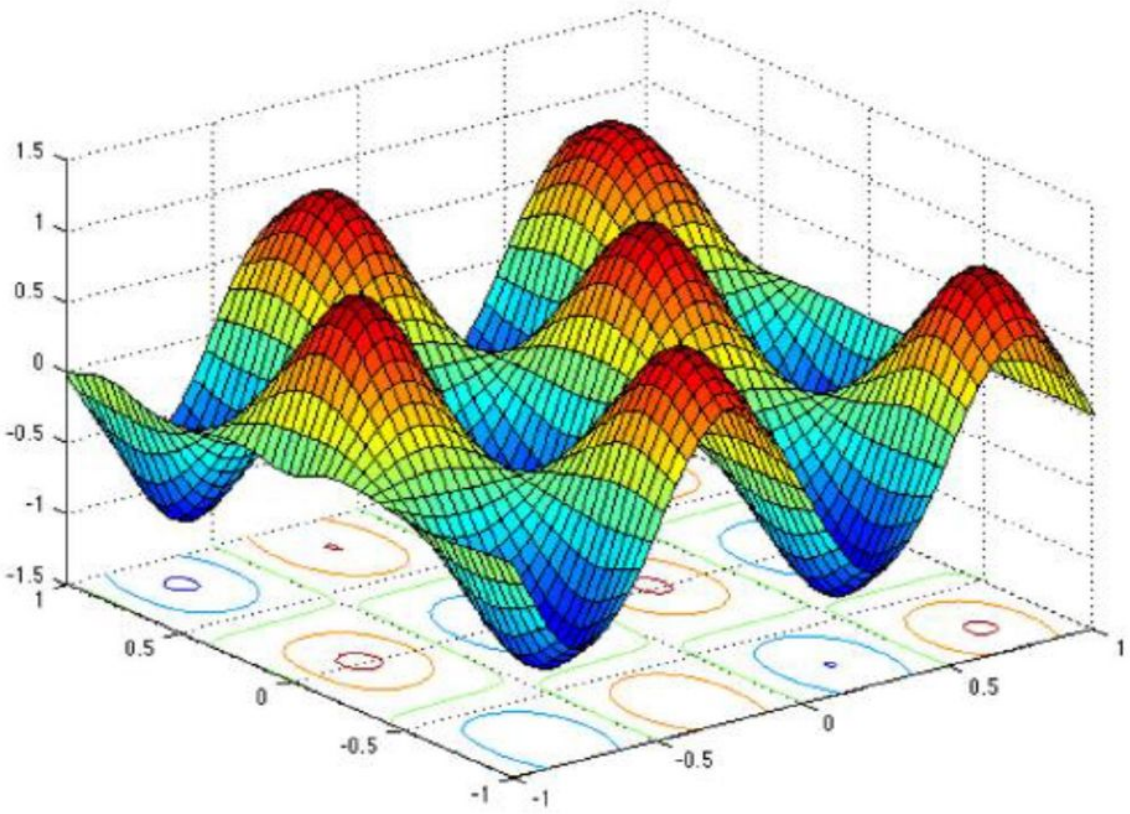

Figure 3

The plot of the piecewise cubic type of MARS model [19]

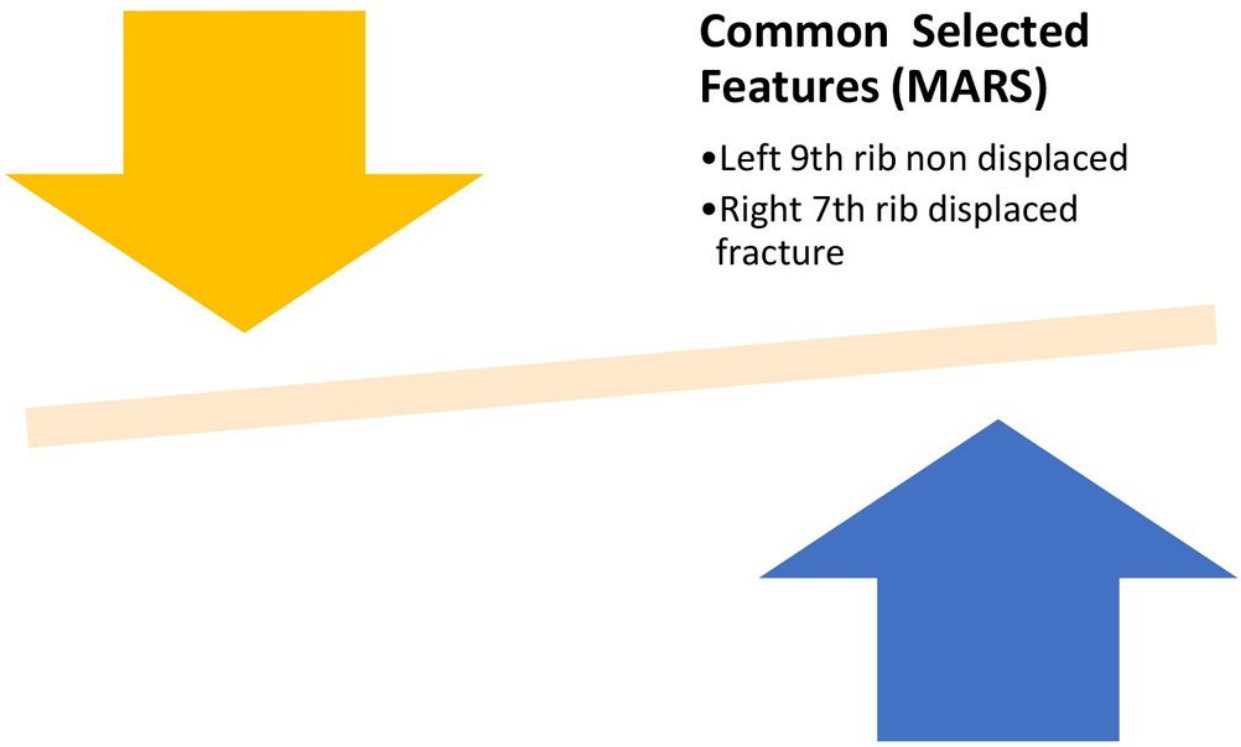

Figure 4

Commonly selected features