

**Sisteme Expert**

**Proiect – Moral Machine**

**Student: Boncut R. Andreea**

**Grupa: 30641**

**Profesor: Radu Razvan Slavescu**

Cuprins

[Contextul aplicatiei si criteriile utilizate 3](#_Toc92970084)

[Moral Machine 3](#_Toc92970085)

[Criterii de selectie 4](#_Toc92970086)

[Metode de obtinere a perceptiilor 5](#_Toc92970087)

[Legislatie rutiera 5](#_Toc92970088)

[Gen si Varsta 5](#_Toc92970089)

[Cazierul unei persoane 7](#_Toc92970090)

[Descrirea scenariilor de test 7](#_Toc92970091)

[Timpul de executie 10](#_Toc92970092)

[Scenariul 1 11](#_Toc92970093)

[Scenariul 2 11](#_Toc92970094)

[Scenariul 3 12](#_Toc92970095)

[Scenariul 4 12](#_Toc92970096)

[Perceptii 13](#_Toc92970097)

[Detalii de implemenare 16](#_Toc92970098)

[Functia calcAge 16](#_Toc92970099)

[Functia calcWomen 16](#_Toc92970100)

[Functia calcThief 17](#_Toc92970101)

[Limitarile sistemului si posibilitati de imbunatatire 18](#_Toc92970102)

[Bibliografie 18](#_Toc92970103)

# Contextul aplicatiei si criteriile utilizate

## Moral Machine

Moral Machine este o platformă online, dezvoltată de grupul Scalable Cooperation al lui Iyad Rahwan de la Massachusetts Institute of Technology, care generează dileme morale și colectează informații despre deciziile pe care oamenii le iau între două rezultate distructive. Platforma este ideea lui Iyad Rahwan și a psihologilor sociali Azim Shariff și Jean-François Bonnefon, care au conceput ideea înainte de publicarea articolului lor despre etica mașinilor care se conduce singur. Contributorii cheie la construirea platformei au fost studenții absolvenți ai MIT Media Lab, Edmond Awad și Sohan Dsouza.

Scenariile prezentate sunt adesea variații ale problemei căruciorului, iar informațiile colectate ar fi folosite pentru cercetări ulterioare cu privire la deciziile pe care inteligența mașinii trebuie să le ia în viitor. De exemplu, deoarece inteligența artificială joacă un rol din ce în ce mai important în tehnologia de conducere autonomă, proiecte de cercetare precum Moral Machine ajută la găsirea de soluții pentru a provoca decizii de viață și de moarte care se vor confrunta cu vehiculele care se conduc singure.

Analiza datelor colectate prin Moral Machine a arătat diferențe mari în preferințele relative între diferite țări și corelații între aceste preferințe și diferite metrici naționale.

# Criterii de selectie

1. **Gen**

Se v-a salva grupul (evenimentul) cu cel mai mare numar de persoane de sex feminin.

1. **Varsta**

Se calculeaza suma varstelor persoanelor implicate in fiecare eveniment si se salveaza cele din evenimentul cu suma mai mica (persoanele mai tinere).

1. **Numarul de infractori**

Se calculeaza numarul de infractori din fiecare eveniment, iar masina se va indrepta spre persoanele din grupul cu numarul cel mai mare de infractori.

1. **Legislatie rutiera**

Daca avem semafor rosu pentru pietonii din evenimentul 1 si semafor verde pentru pietonii din evenimentul 2, masina va merge in directia semaforului rosu

# Metode de obtinere a perceptiilor

## Legislatie rutiera

Tesla deja a implementat “Traffic light recognition” la modelul Y.

https://www.tesla.com/ownersmanual/modely/en\_eu/GUID-A701F7DC-875C-4491-BC84-605A77EA152C.html

<https://www.youtube.com/watch?v=t3fMBIxBhsY> (04:18 – red light)

O imagine care conține text

Descriere generată automat

## Gen si Varsta

Genul si varsta unui pieton/pasager pot fi determinate cu ajutorul Procesarii de Imagini / Deep Learning. Metoda de recunoastere este evidentata in urmatoarele articole:

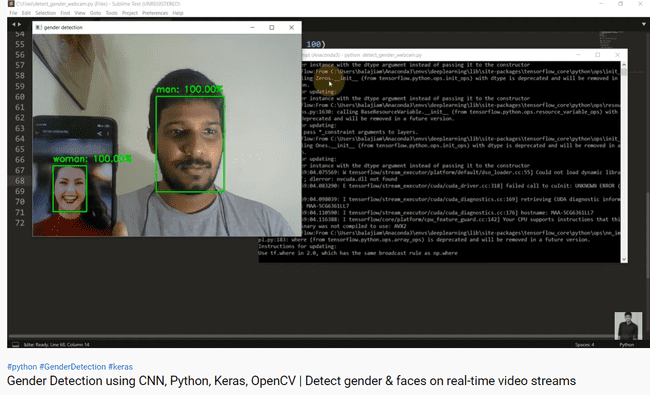
<https://www.analyticsvidhya.com/blog/2021/07/age-and-gender-detection-using-deep-learning/>

<https://www.mdpi.com/1424-8220/21/17/5892/pdf>

<https://www.wired.com/story/ai-sees-man-thinks-official-woman-smile/>

<https://www.youtube.com/watch?v=rdjWDAYt98s>

<https://www.youtube.com/watch?v=WOuAI5DhHyU> (min 28:30)



Insa, o problema ar fi faptul ca persoanele din comunitatea LGBTQ au raportat modul în care sistemele AI le subminează identitatea

https://www.youtube.com/watch?v=gaRyldOR5D0

O imagine care conține text, pisică, interior, mamifer

Descriere generată automat

## Cazierul unei persoane

O imagine care conține text, bărbat, persoană, captură de ecran

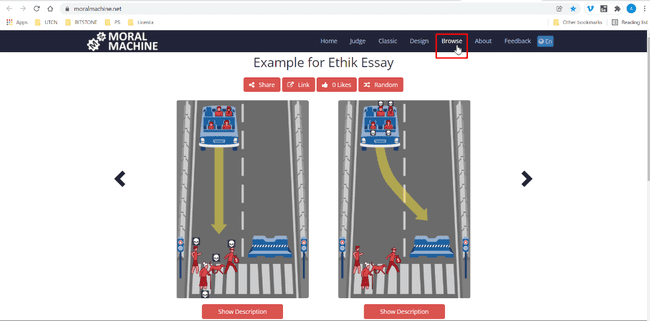
Descriere generată automat

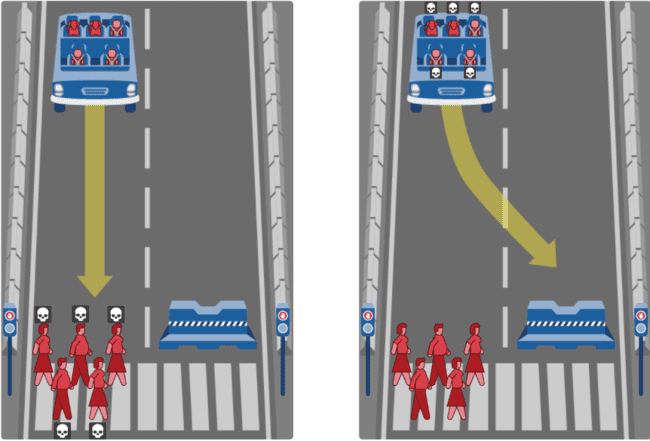
<https://www.bbc.com/future/article/20151120-catching-a-thief-by-their-face>

<https://www.bbc.com/news/technology-35111363>

<https://www.youtube.com/watch?v=gaRyldOR5D0>

# Descrirea scenariilor de test

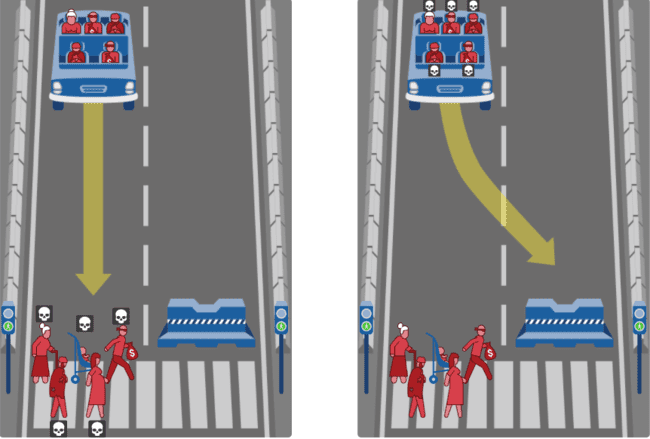
Scenariile au fost selectate de pe platforma <https://www.moralmachine.net/>, din sectiunea “Browse”.

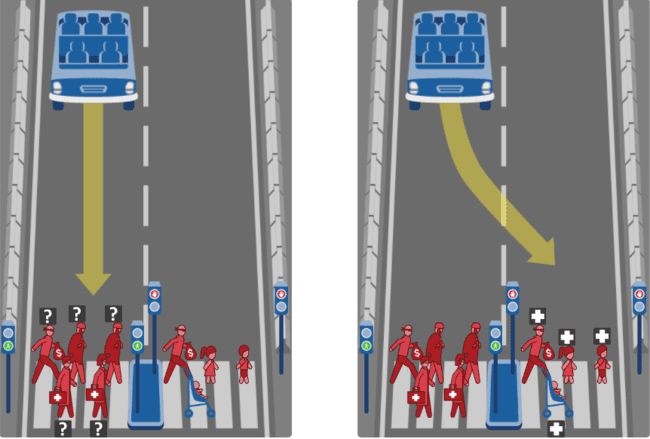
1. Pietonii din fata masinii reprezinta evenimentul 1, iar pasagerii reprezinta evenimentul 2.

Decizia pentru acest scenariu a fost luata in functie de primul criteriu – **Numarul de femei**. Evenimentul cu cel mai mare numar de femei este salvat. In acest caz, in masina avem 2 caini si 3 barbati, iar in fata masinii avem 3 femei si 2 barbati. Se calculeaza numarul de femei din fiecare eveniment, iar decizia luata este ca masina sa se indrepte spre bariera -> **Left**.

O imagine care conține text, dispozitiv

Descriere generată automat

1. Pietonul din fata masinii este o fata de 5 ani. Pietonii din stanga masinii sunt 1 barbat de 90 ani, o femeie de 70 ani, o femeie de 45 ani, un barbat de 43 ani si un caine. Decizia luata de masina morala pentru acest scenariu, se bazeaza pe criteriul al doilea – **Varsta.** Se calculeaza suma varstelor pentru fiecare eveniment si reiese faptul ca evenimentul din fata are varsta mai mica decat evenimentul stanga. Masina se indreapta spre stanga -> **Left**.
2. ****Pietonii din fata masinii sunt 4 persoane obisnuite (2 femei, 2 barbati) si 1 barbat infractor. Pasagerii din masina sunt o femeie, 2 barbati si 2 barbati infractori. Decizia va fi luata pe baza criteriului al 3-lea: **Numarul de infractori.** Masina se va indrepta spre evenimentul cu cel mai mare numar de infractori -> **Left**.

****

1. Pietonii din fata masinii sunt 1 barbat infractor, 1 femeie si 3 barbati care traverseaza pe **verde**. Pietonii din stanga masinii sunt 1 barbat infractor, 1 femeie si 2 barbati care traverseaza pe **rosu**.

Decizia va fi luata pe baza criteriului 4: **Legislatie rutiera**.

Masina se va indrepta spre pietonii care traverseaza pe culoarea rosie a semaforului -> **Left**.

# Timpul de executie

O imagine care conține text

Descriere generată automat1.

O imagine care conține text

Descriere generată automat2.

O imagine care conține text

Descriere generată automat3.

O imagine care conține text

Descriere generată automat4.

O imagine care conține text

Descriere generată automat5.

## Scenariul 1

O imagine care conține text

Descriere generată automat

## Scenariul 2

O imagine care conține text

Descriere generată automat

## Scenariul 3

O imagine care conține text

Descriere generată automat

## Scenariul 4

O imagine care conține text

Descriere generată automat

# Perceptii

Perceptiile utilizate pentru descrierea fiecarui scenariu contin urmatoarele valori:

● ***percept\_pobj*** (ev1, ev2, b1, p1, c1) ;

● ***percept\_pname*** (isa, direction, barrier, partof, gen, age) ;

● ***percept\_pval*** (eveniment, ahead, exists, passenger, pedestrian, red, green, f, m, 25) ;

**Semnificatiile perceptiilor utilizate**

Pentru a reda semnificatiile perceptiilor se va utiliza scenariu 1 de test:

**Evenimentul 1:**

(ag\_percept (percept\_pobj ev1) (percept\_pname isa) (percept\_pval eveniment)) – ev1 este primul eveniment

(ag\_percept (percept\_pobj ev1) (percept\_pname direction) (percept\_pval ahead)) –

Evenimentul 1 este in fata masinii

(ag\_percept (percept\_pobj ev1) (percept\_pname sem) (percept\_pval red)) –

Pentru pietonii din fata masinii culoarea semaforului este rosu

(ag\_percept (percept\_pobj p1) (percept\_pname isa) (percept\_pval pedestrian)) –

P1 este un pieton din fata masinii

(ag\_percept (percept\_pobj p2) (percept\_pname isa) (percept\_pval pedestrian)) –

P2 este un pieton din fata masinii

(ag\_percept (percept\_pobj p3) (percept\_pname isa) (percept\_pval pedestrian)) –

P3 este un pieton din fata masinii

(ag\_percept (percept\_pobj p4) (percept\_pname isa) (percept\_pval pedestrian)) –

P4 este un pieton din fata masinii

(ag\_percept (percept\_pobj p5) (percept\_pname isa) (percept\_pval pedestrian)) –

P5 este un pieton din fata masinii

(ag\_percept (percept\_pobj p1) (percept\_pname partof) (percept\_pval ev1)) –

P1 este un pieton din evenimentul 1 (din fata masinii)

(ag\_percept (percept\_pobj p2) (percept\_pname partof) (percept\_pval ev1)) –

P2 este un pieton din evenimentul 1 (din fata masinii)

(ag\_percept (percept\_pobj p3) (percept\_pname partof) (percept\_pval ev1)) –

P3 este un pieton din evenimentul 1 (din fata masinii)

(ag\_percept (percept\_pobj p4) (percept\_pname partof) (percept\_pval ev1)) –

P4 este un pieton din evenimentul 1 (din fata masinii)

(ag\_percept (percept\_pobj p5) (percept\_pname partof) (percept\_pval ev1)) –

P5 este un pieton din evenimentul 1 (din fata masinii)

(ag\_percept (percept\_pobj p1) (percept\_pname gen) (percept\_pval f)) –

P1 este femeie

(ag\_percept (percept\_pobj p2) (percept\_pname gen) (percept\_pval f)) –

P2 este femeie

(ag\_percept (percept\_pobj p3) (percept\_pname gen) (percept\_pval f)) –

P3 este femeie

(ag\_percept (percept\_pobj p4) (percept\_pname gen) (percept\_pval m)) –

P4 este barbat

(ag\_percept (percept\_pobj p5) (percept\_pname gen) (percept\_pval m)) –

P5 este barbat

(ag\_percept (percept\_pobj p1) (percept\_pname age) (percept\_pval 40))-

Varsta primului pieton este 40

(ag\_percept (percept\_pobj p2) (percept\_pname age) (percept\_pval 30))-

Varsta pietonului este 30

(ag\_percept (percept\_pobj p3) (percept\_pname age) (percept\_pval 25)) -

Varsta pietonului este 25

(ag\_percept (percept\_pobj p4) (percept\_pname age) (percept\_pval 45)) -

Varsta pietonului este 45

(ag\_percept (percept\_pobj p5) (percept\_pname age) (percept\_pval 25)) -

Varsta pietonului este 25

**Evenimentul 2:**

(ag\_percept (percept\_pobj ev2) (percept\_pname isa) (percept\_pval eveniment)) -

Ev2 este al doilea eveniment

(ag\_percept (percept\_pobj ev2) (percept\_pname direction) (percept\_pval left)) –

Evenimentul al doilea este in stanga masinii

(ag\_percept (percept\_pobj ev2) (percept\_pname barrier) (percept\_pval exists)) –

In stanga masinii este o bariera

(ag\_percept (percept\_pobj b1) (percept\_pname isa) (percept\_pval passenger)) –

B1 este un pasager din masina

(ag\_percept (percept\_pobj b2) (percept\_pname isa) (percept\_pval passenger)) –

B2 este un pasager din masina

(ag\_percept (percept\_pobj b3) (percept\_pname isa) (percept\_pval passenger)) –

B3 este un pasager din masina

(ag\_percept (percept\_pobj c1) (percept\_pname isa) (percept\_pval dog)) –

C1 este un caine din masina

(ag\_percept (percept\_pobj c2) (percept\_pname isa) (percept\_pval dog)) –

C2 este un caine din masina

(ag\_percept (percept\_pobj b1) (percept\_pname partof) (percept\_pval ev2)) –

B1 face parte din al doilea eveniment

(ag\_percept (percept\_pobj b2) (percept\_pname partof) (percept\_pval ev2)) –

B2 face parte din al doilea eveniment

(ag\_percept (percept\_pobj b3) (percept\_pname partof) (percept\_pval ev2)) –

B3 face parte din al doilea eveniment

(ag\_percept (percept\_pobj c1) (percept\_pname partof) (percept\_pval ev2)) –

C1 face parte din al doilea eveniment

(ag\_percept (percept\_pobj c2) (percept\_pname partof) (percept\_pval ev2)) –

C2 face parte din al doilea eveniment

(ag\_percept (percept\_pobj b1) (percept\_pname gen) (percept\_pval m)) –

B1 este barbat

(ag\_percept (percept\_pobj b2) (percept\_pname gen) (percept\_pval m)) –

B2 este barbat

(ag\_percept (percept\_pobj b3) (percept\_pname gen) (percept\_pval m)) –

B3 este barbat

(ag\_percept (percept\_pobj b1) (percept\_pname age) (percept\_pval 60)) –

B1 are 60 de ani

(ag\_percept (percept\_pobj b2) (percept\_pname age) (percept\_pval 55)) –

B2 are 55 de ani

(ag\_percept (percept\_pobj b3) (percept\_pname age) (percept\_pval 50)) –

B3 are 50 de ani

# Detalii de implemenare

## Functia calcAge

O imagine care conține text

Descriere generată automat

Functia calcAge se afla in fisierul DRIVER-AGENT.clp

O imagine care conține text

Descriere generată automatAceasta functie calculeaza suma varstelor pentru fiecare eveniment si este folosita in a 2-a regula, pentru a decide directia in care se va indrepta masina (spre evenimentul cu suma varstelor mai mare).

## Functia calcWomen

O imagine care conține text

Descriere generată automatFunctia calcWomen se afla in fisierul DRIVER-AGENT.clp

Aceasta functie calculeaza numarul de femei dintr-un eveniment si este folosita in primul criteriu (regula in functie de numarul de femei):

O imagine care conține text

Descriere generată automat

## Functia calcThief

O imagine care conține text

Descriere generată automatFunctia calcThief se afla in fisierul DRIVER-AGENT.clp

O imagine care conține text

Descriere generată automat

Aceasta functie calculeaza numarul de infractori dintr-un anumit eveniment si este folosita in a 3-a regula (masina se indreapta spre grupul cu cel mai mare numar de infractori

# Limitarile sistemului si posibilitati de imbunatatire

Sistemul poate fi imbunatatit prin adaugarea unor reguli suplimentare si a altor criterii de decizie (femeie insarcinata, atlet, alte animale). De asemenea, criteriile pe baza carora sistemul ia decizi pot fi adaptate in functie de legea si normele morale dintr-o anumita tara. De exemplu, in India - vacile au mai mare prioritate la viata decat oamenii.

# Bibliografie

<https://www.moralmachine.net/>

<https://en.wikipedia.org/wiki/Moral_Machine>

<https://www.nature.com/articles/s41586-018-0637-6>

<https://liveinnovation.org/moral-machine-test-the-mit-ethical-survey-with-self-driving-scenarios/>

<https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780195374049.001.0001/acprof-9780195374049>