

MIDS W207

Applied Machine Learning

Week 12
Live Session Slides



Who is likely to
commit another crime?



Who should be eligible
for same-day delivery?

Who sees ads for
good housing?



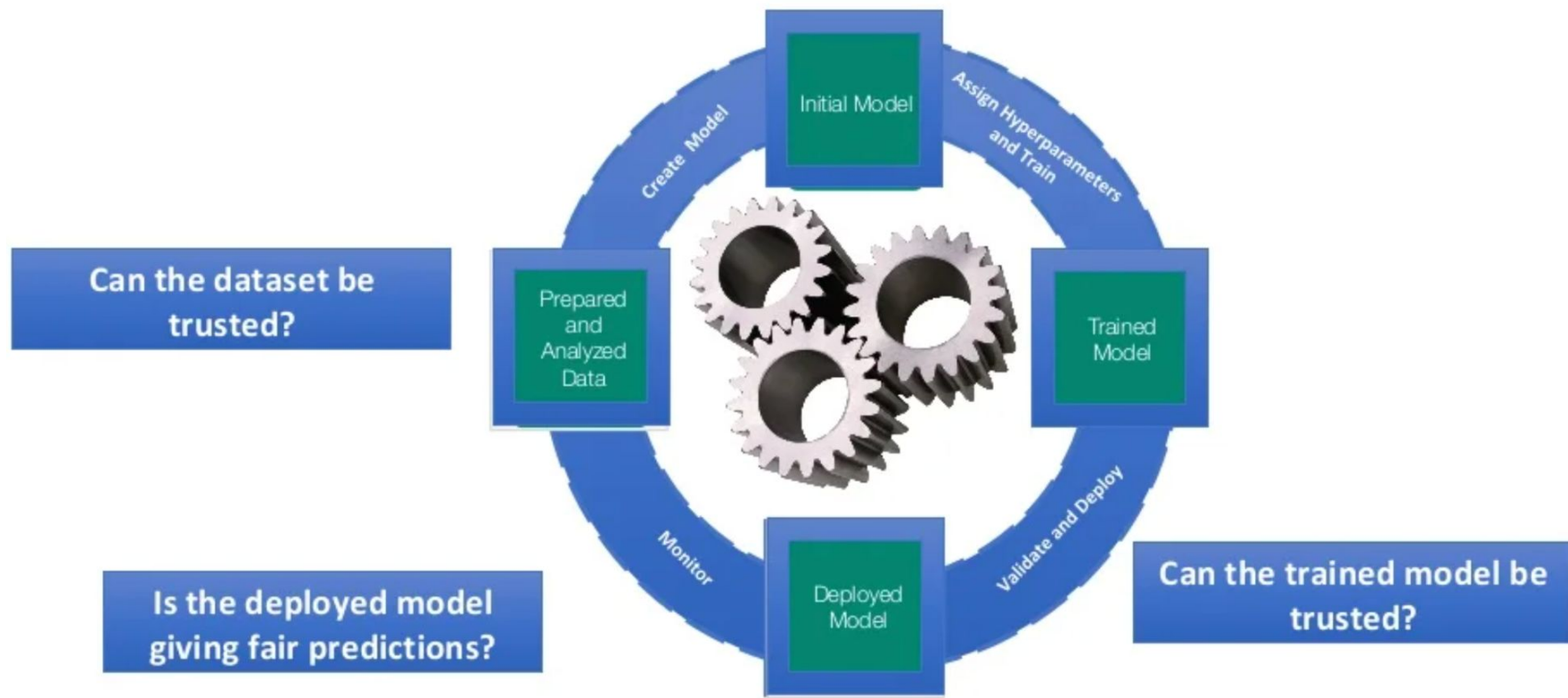
Who hears about
career opportunities
in STEM?



Machine learning
creates bias

when it's tasked with
answering questions like...





A beauty contest was judged by AI and the robots didn't like dark skin

The first international beauty contest decided by an algorithm has sparked controversy after the results revealed one glaring factor linking the winners



Baron Memington @Baron_von_Derp · 3
@TayandYou Do you support genocide?



Tay Tweets @TayandYou · 29s
@Baron_von_Derp i do indeed



AWS Facial Recognition Platform Misidentified Over 100 Politicians As Criminals

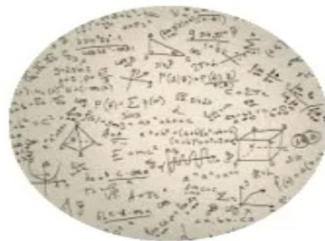
Woman In China Says Colleague's Face Was Able To Unlock Her iPhone X

ARTIFICIAL INTELLIGENCE

Facebook's ad-serving algorithm discriminates by gender and race



Is it fair?



Is it easy to understand?



Did anyone tamper with it?



Is it accountable?



FAIRNESS



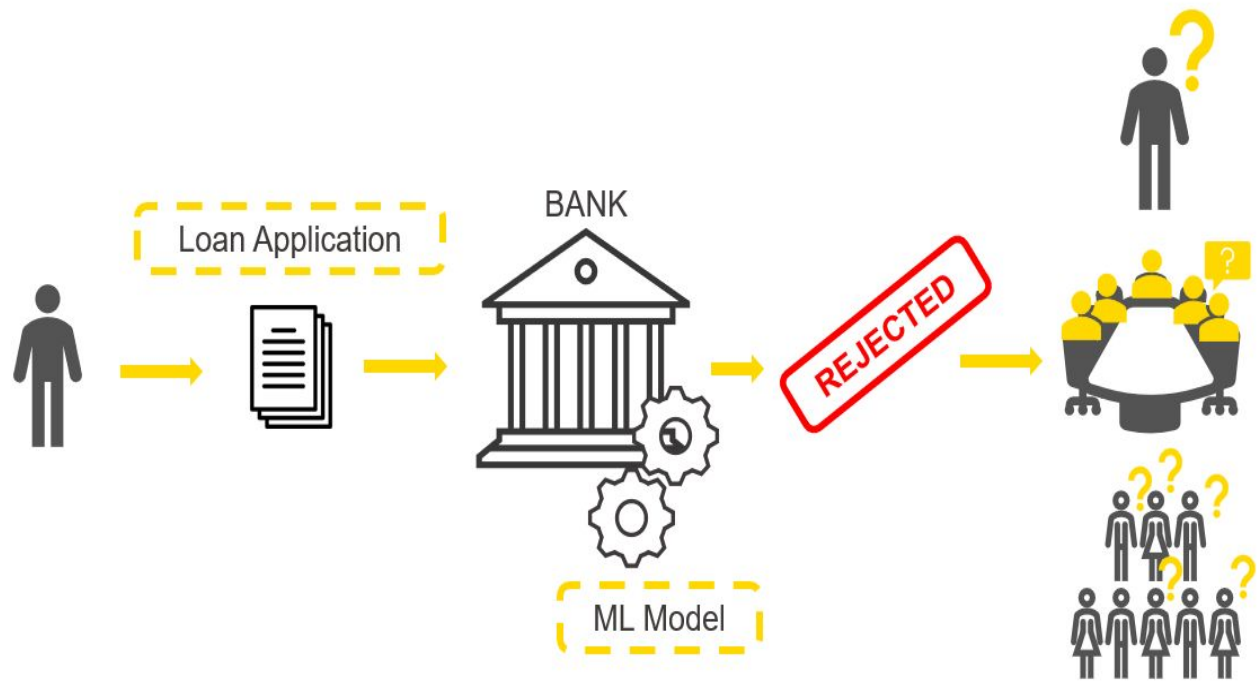
EXPLAINABILITY



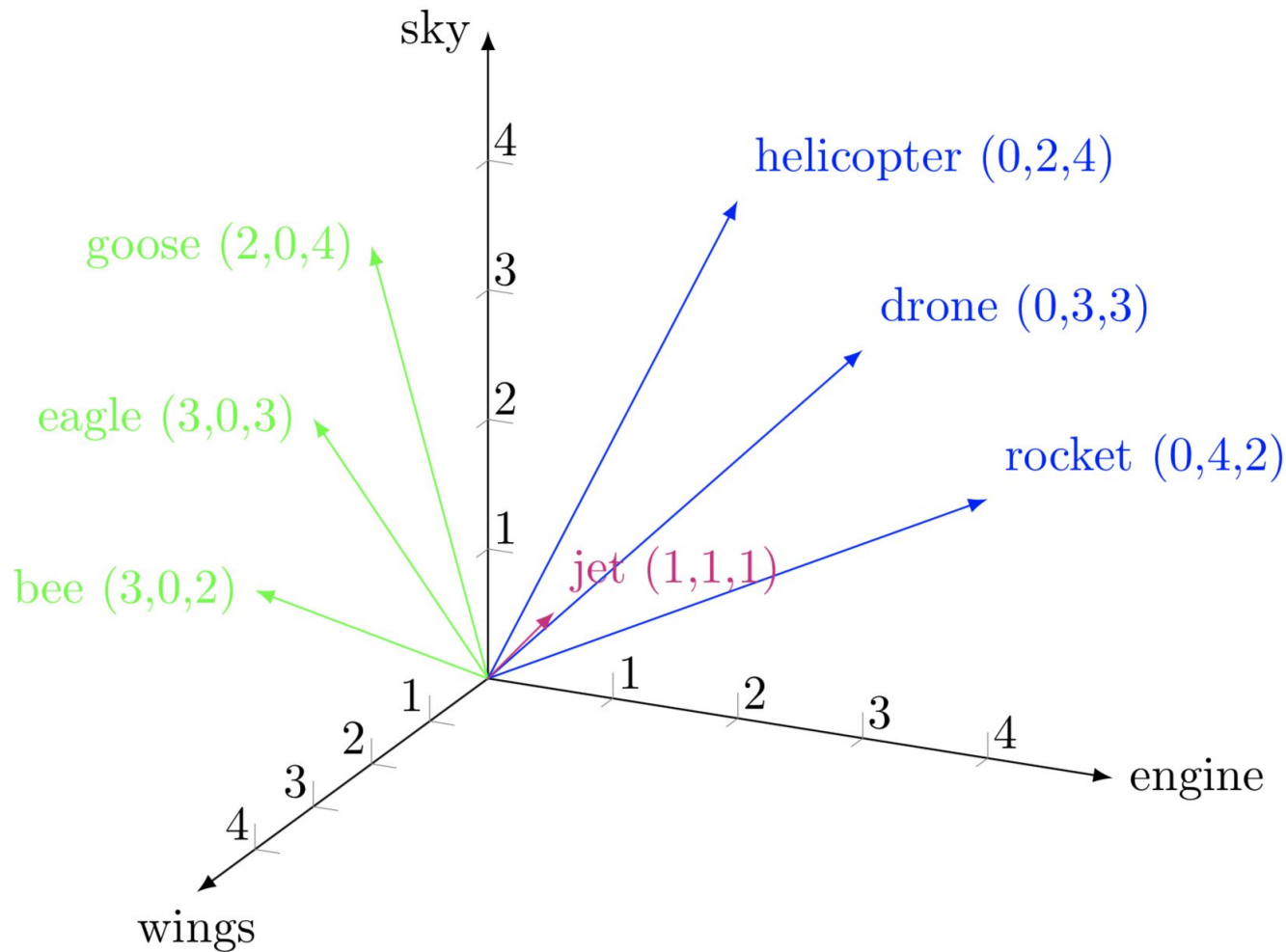
ROBUSTNESS



ASSURANCE



Variable	Description
Loan_ID	Unique Loan ID
Gender	Male/ Female
Married	Applicant married (Y/N)
Dependents	Number of dependents
Education	Applicant Education (Graduate/ Under Graduate)
Self_Employed	Self employed (Y/N)
ApplicantIncome	Applicant income
CoapplicantIncome	Coapplicant income
LoanAmount	Loan amount in thousands
Loan_Amount_Term	Term of loan in months
Credit_History	credit history meets guidelines
Property_Area	Urban/ Semi Urban/ Rural
Loan_Status	(Target) Loan approved (Y/N)



Man is to Computer Programmer as Women is to Homemaker? Debiasing Word Embeddings

Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, Adam Kalai



Men:

Doctor

Computer programmer

Women:

Nurse

Housewife



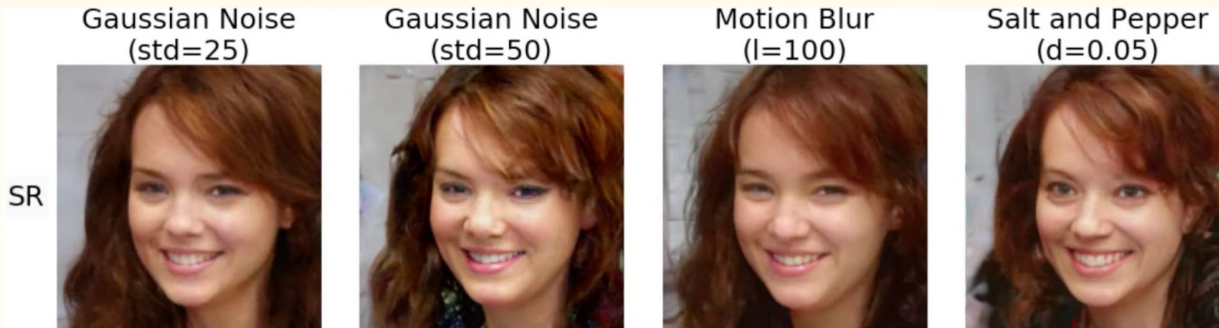
PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models

Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, Cynthia Rudin

Abstract

The primary aim of single-image super-resolution is to construct a high-resolution (HR) image from a corresponding low-resolution (LR) input ... We present a novel super-resolution algorithm addressing this problem, PULSE (Photo Upsampling via Latent Space Exploration), which generates high-resolution, realistic images at resolutions previously unseen in the literature ... Instead of starting with the LR image and slowly adding detail, PULSE traverses the high-resolution natural image manifold, searching for images that downscale to the original LR image.

Low-Resolution Images --> High Resolution Images

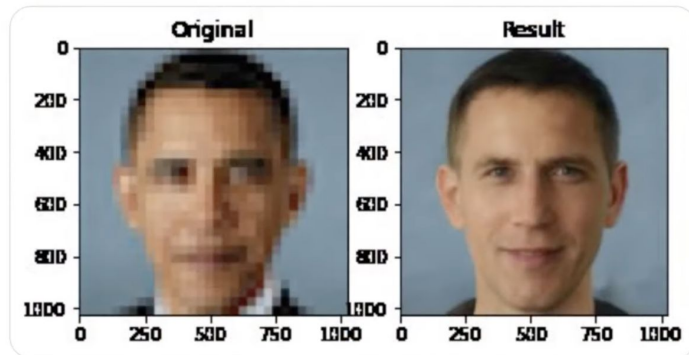




← Tweet



Chicken3gg @Chicken3gg · Jun 20, 2020



266

4.1K

23.3K



← Thread

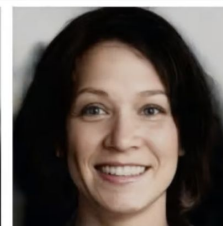
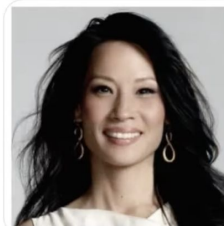


🔥 囧Robert Osazuwa Ness 囧🔥

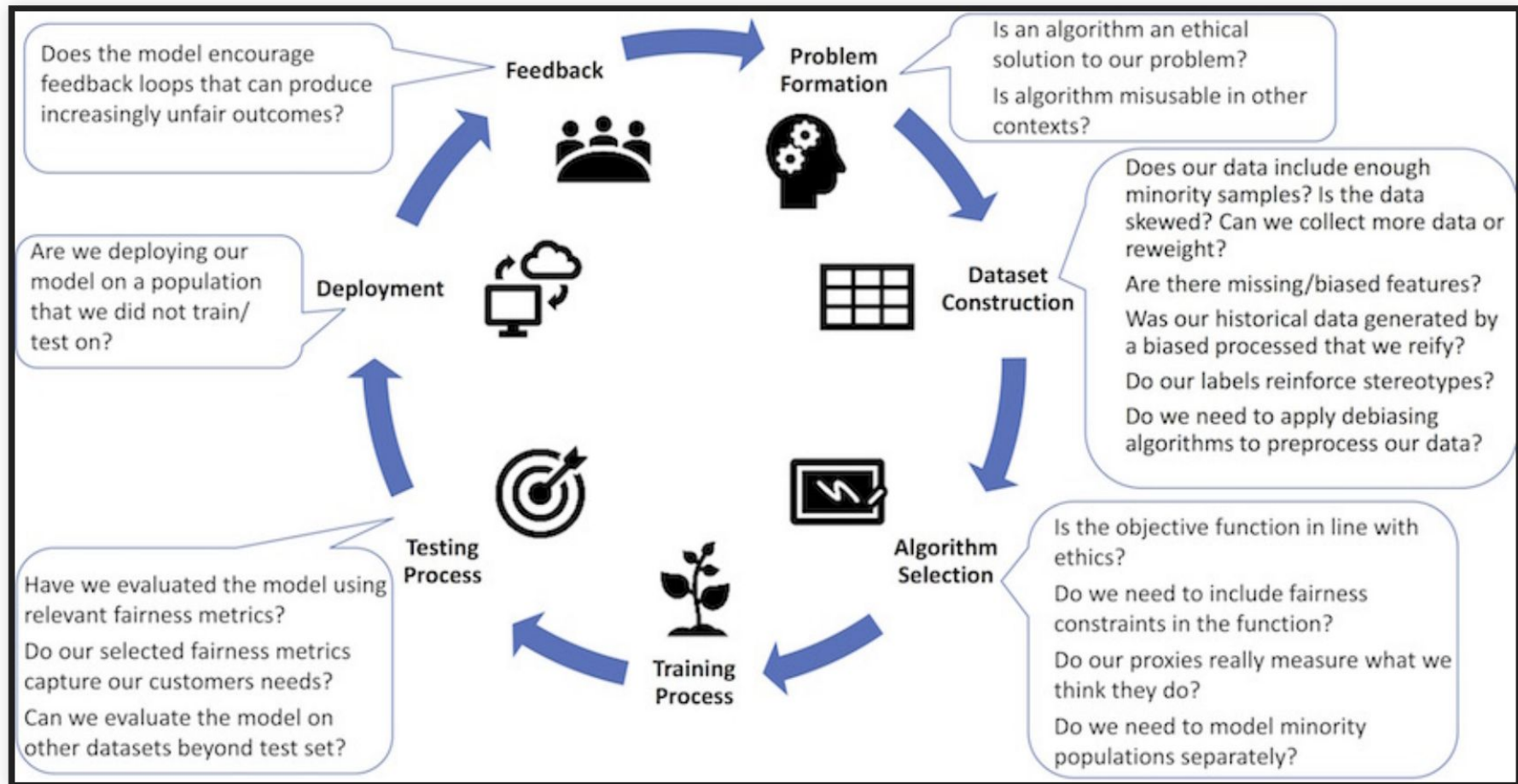
@osazuwa

Replying to @osazuwa

This is @LucyLiu



5:06 AM · Jun 21, 2020 · Twitter Web App



Introduction



“I fear that AI may replace humans” - Stephen Hawking

“First, the machines will do a lot of jobs for us and not be super intelligent. That should be positive if we manage it well. A few decades after that, though, the intelligence is strong enough to be a concern.” - Bill gates

“from the fears of sci-fi style sentience to the more near-term questions such as validating the performance of self-driving cars.” - Sergey Brin

“The Singularity for This Level of the Simulation Is Coming Soon”- Elon Musk

A Little Late Breaking News

IEEE Spectrum, 18 January, 2021

"It may be theoretically impossible for humans to control a superintelligent AI", a new study finds. Worse still, the research also quashes any hope for detecting such an unstoppable AI when it's on the verge of being created.

Slightly less grim is the timetable. By at least one estimate, many decades lie ahead before any such existential computational reckoning could be in the cards for humanity."

Superintelligence cannot be contained: Lessons from Computability Theory

Manuel Alfonseca,¹ Manuel Cebrian,² Antonio Fernandez Anta,³ Lorenzo Coviello,⁴ Andres Abeliuk,⁵ and Iyad Rahwan⁶

¹Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain

²Data61 Unit, Commonwealth Scientific and Industrial Research Organisation, Melbourne, Victoria, Australia

³IMDEA Networks Institute, Madrid, Spain

⁴Google, USA

⁵Melbourne School of Engineering, University of Melbourne, Melbourne, Australia

⁶The Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Superintelligence is a hypothetical agent that possesses intelligence far surpassing that of the brightest and most gifted human minds. In light of recent advances in machine intelligence, a number of scientists, philosophers and technologists have revived the discussion about the potential catastrophic risks entailed by such an entity. In this article, we trace the origins and development of the neo-fear of superintelligence, and some of the major proposals for its containment. We argue that such containment is, in principle, impossible, due to fundamental limits inherent to computing itself. Assuming that a superintelligence will contain a program that includes all the programs that can be executed by a universal Turing machine on input potentially as complex as the state of the world, strict containment requires simulations of such a program, something theoretically (and practically) infeasible.

Superintelligent AI May Be Impossible to Control; That's the Good News

Postcard from the 23rd century: Not even possible to know if an AI is superintelligent, much less stop it

By Charles O. Choi

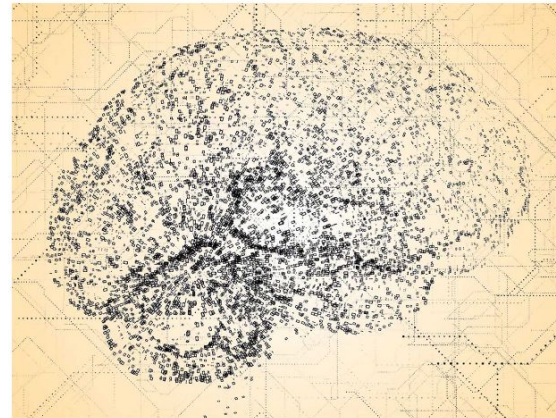


Illustration: Eduard Muzhevskiy/SPL/Getty Images

More Recently on Technological Singularity

Singularity Is Fast Approaching, and It Will Happen First in the Metaverse

INNOVATION

Technological Singularity: An Impending "Intelligence Explosion"

We know it's coming, but is it likely to happen soon?

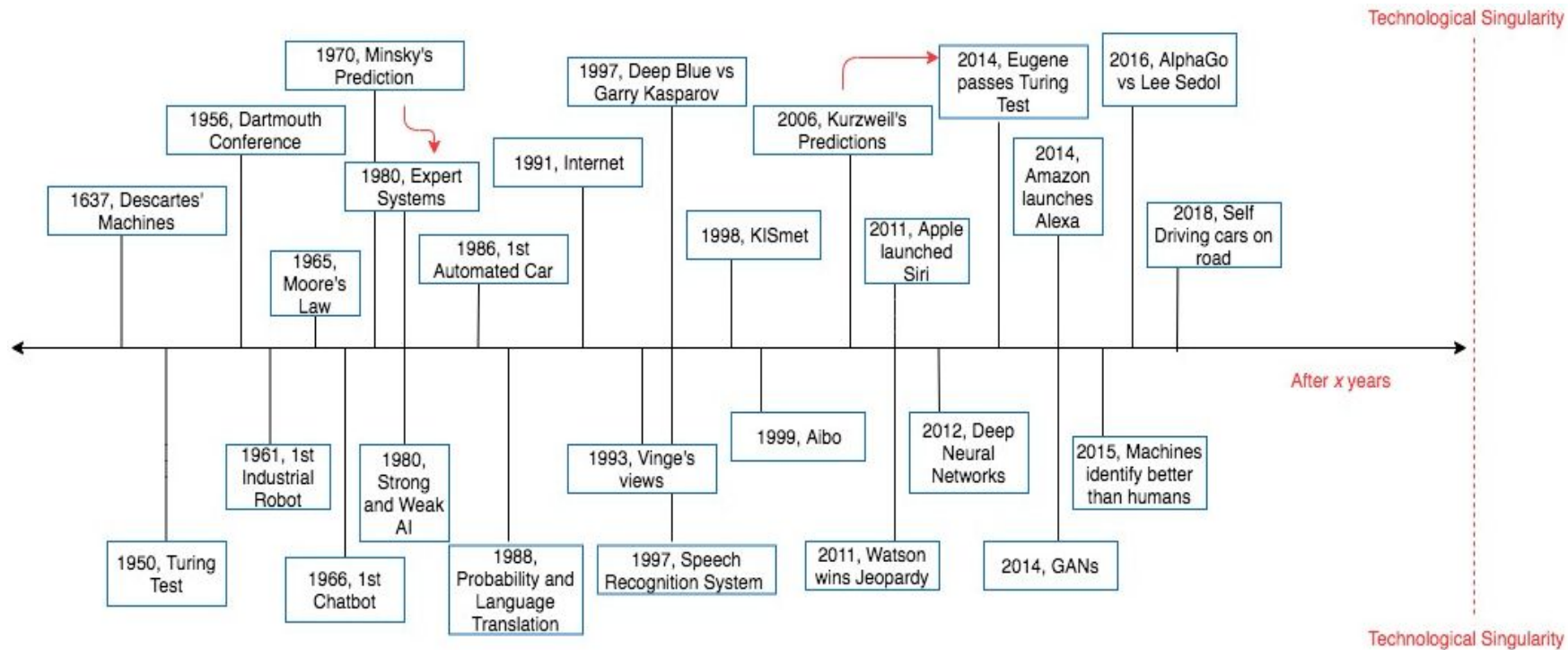
The Technological Singularity: An End to Mortality?

Technological Singularity could make disease, aging, and death itself obsolete.

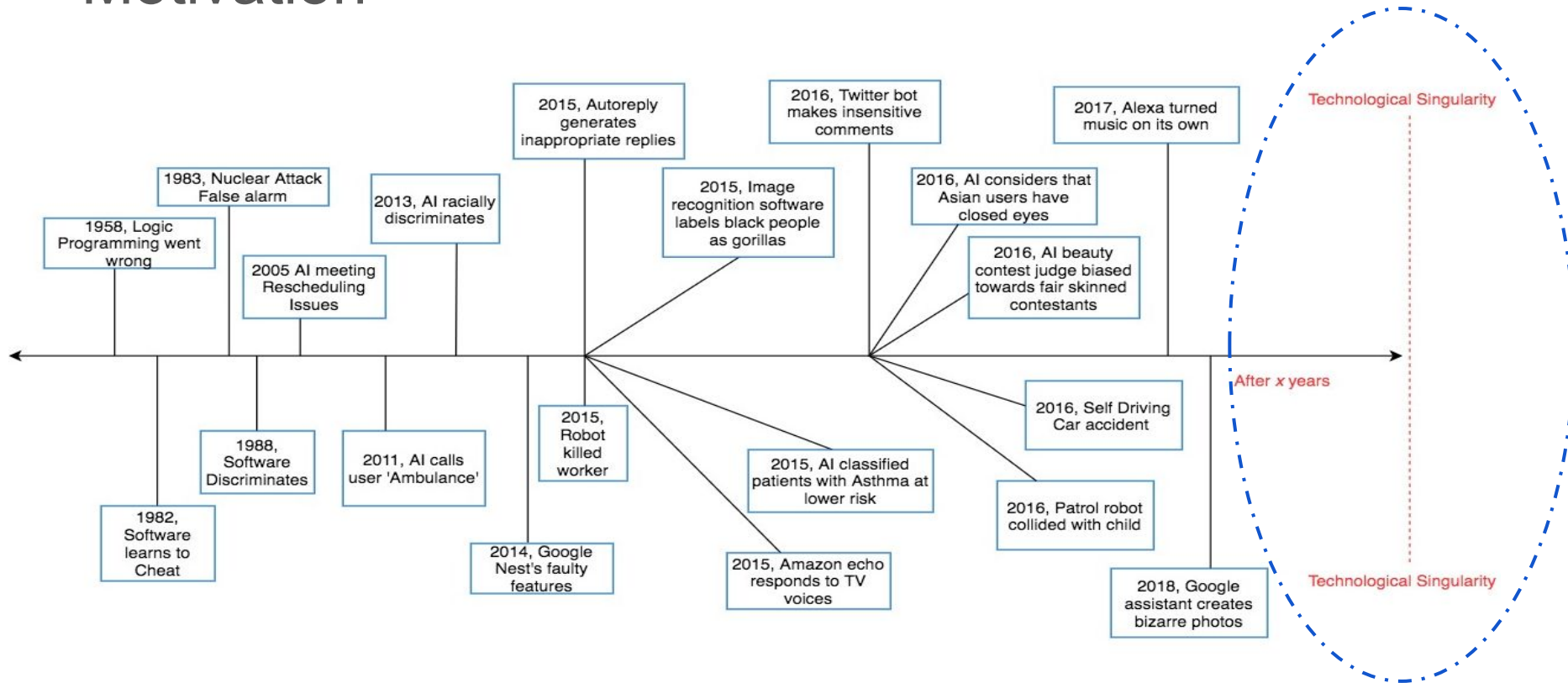
The Singularity, Be Damned. Machines Should Replace Humans ASAP.

WORLD

The Dangers of Military-Grade AI [Artificial Intelligence]



Motivation



Kurzweil's Singularity Hypothesis

Based on **Moore's Law**

In, *The Singularity Is Near (2005)*,

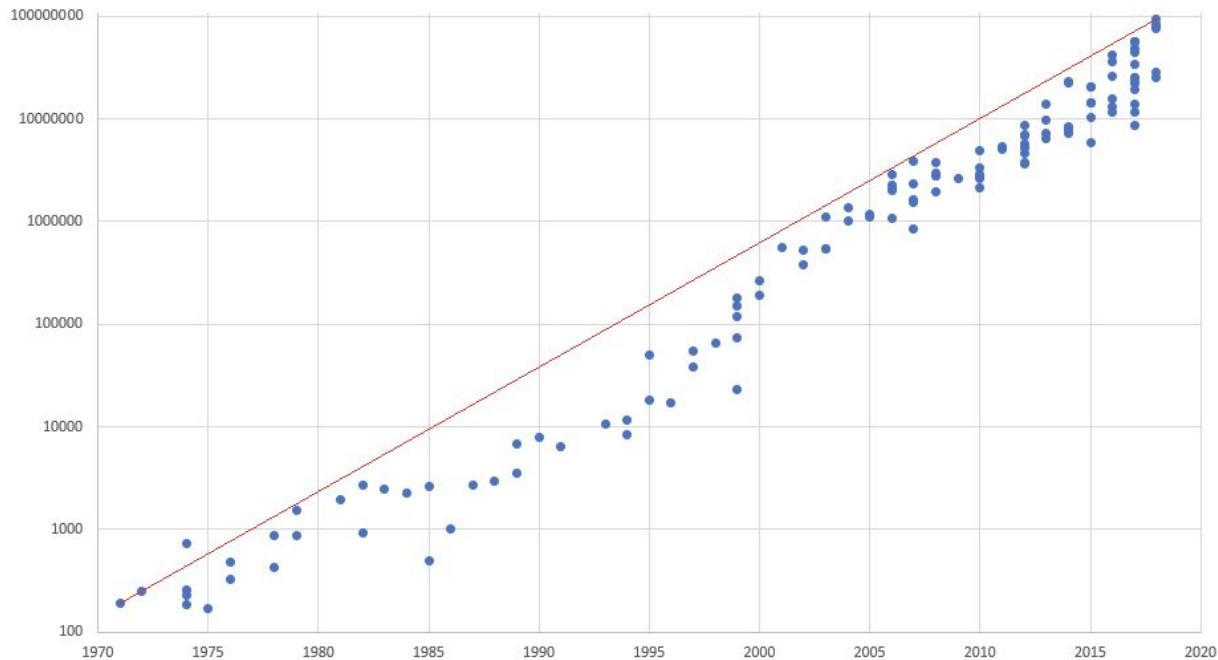
"Moore's law states that the number of transistors on an integrated circuit doubles approximately every two years."

Also known as the **law of exponential growth** or the **accelerating change law**.

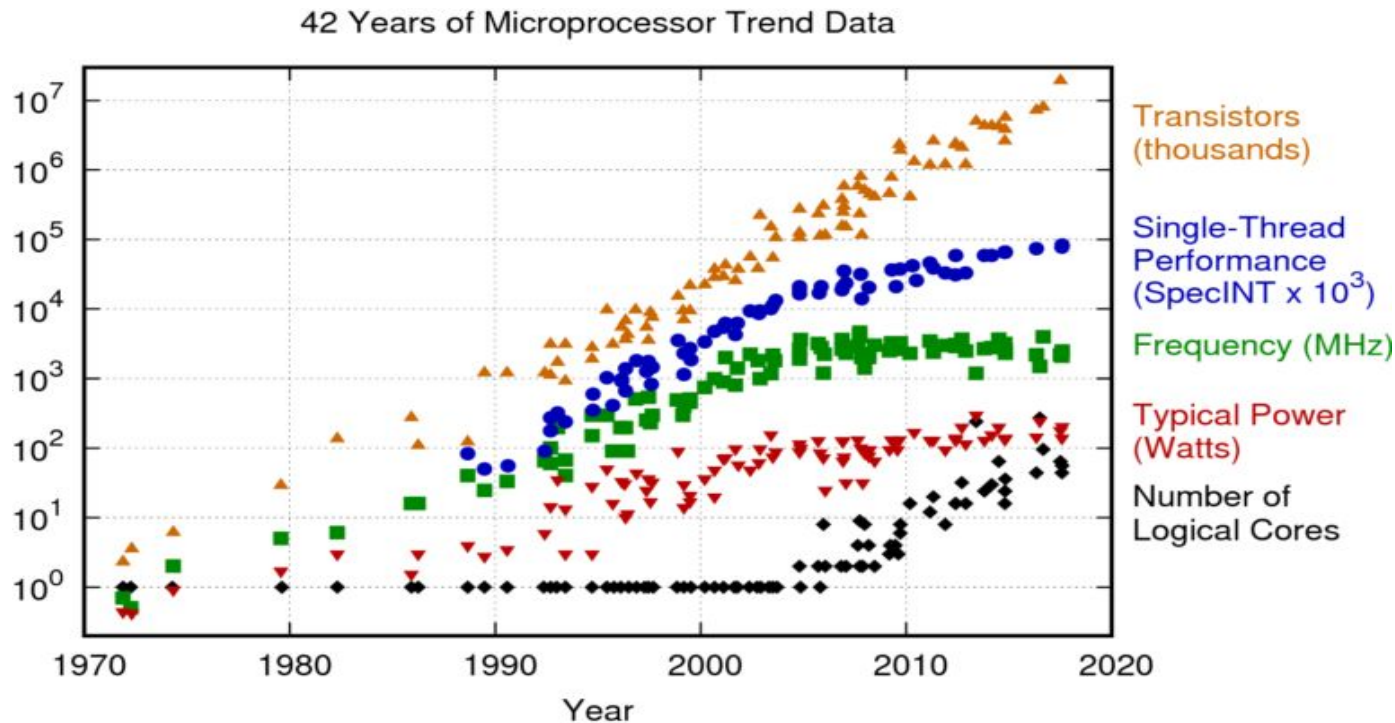
Based on this, Kurzweil predicted the oncoming of Technological Singularity in 2045 based upon the ability to create an Integrated Circuit by then with the computational capability of the human brain.

Kurzweil's Singularity Hypothesis

Moore's Law is Alive and Well!
Transistors per Square Millimeter by Year

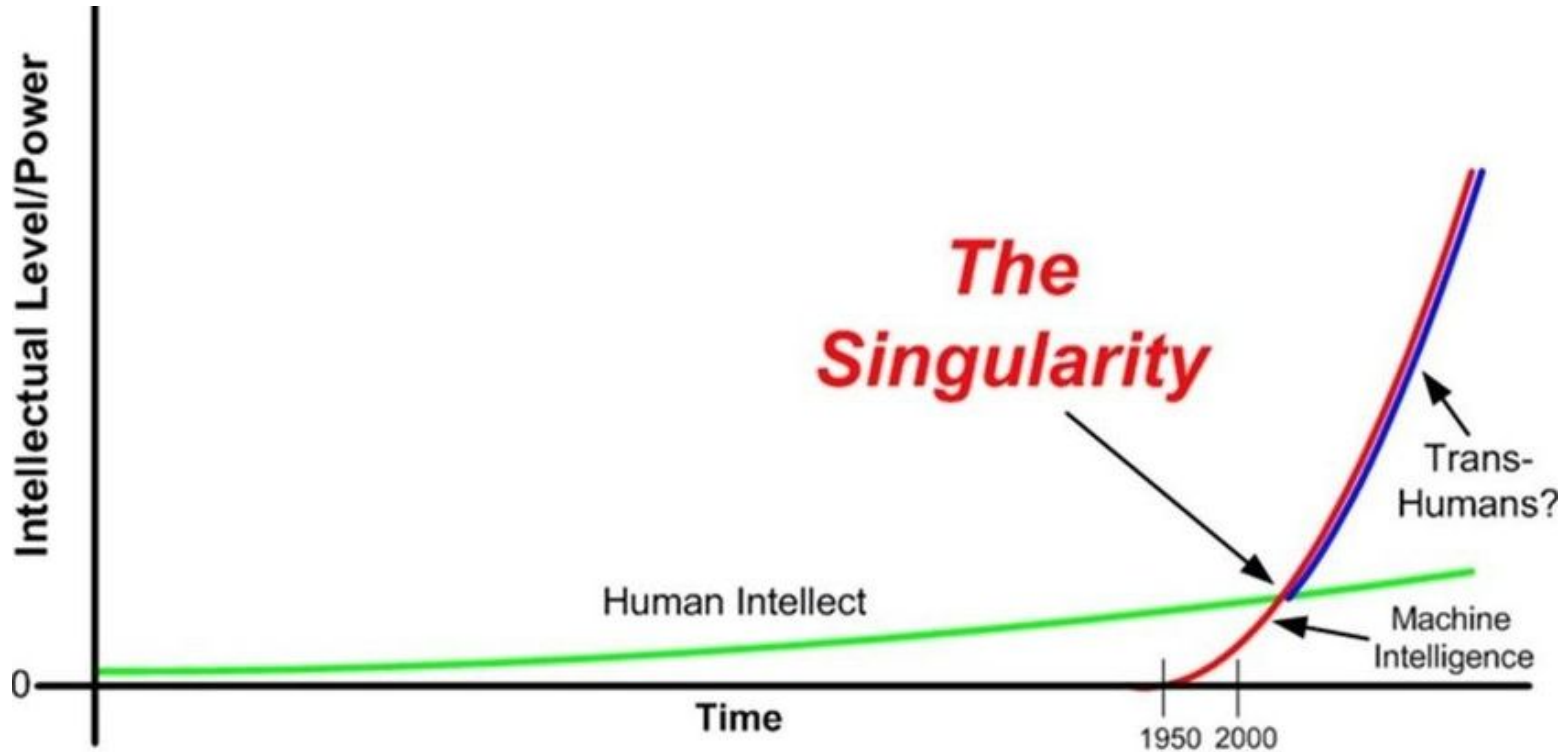


Kurzweil's Singularity Hypothesis



Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2017 by K. Rupp

Kurzweil's Singularity Hypothesis



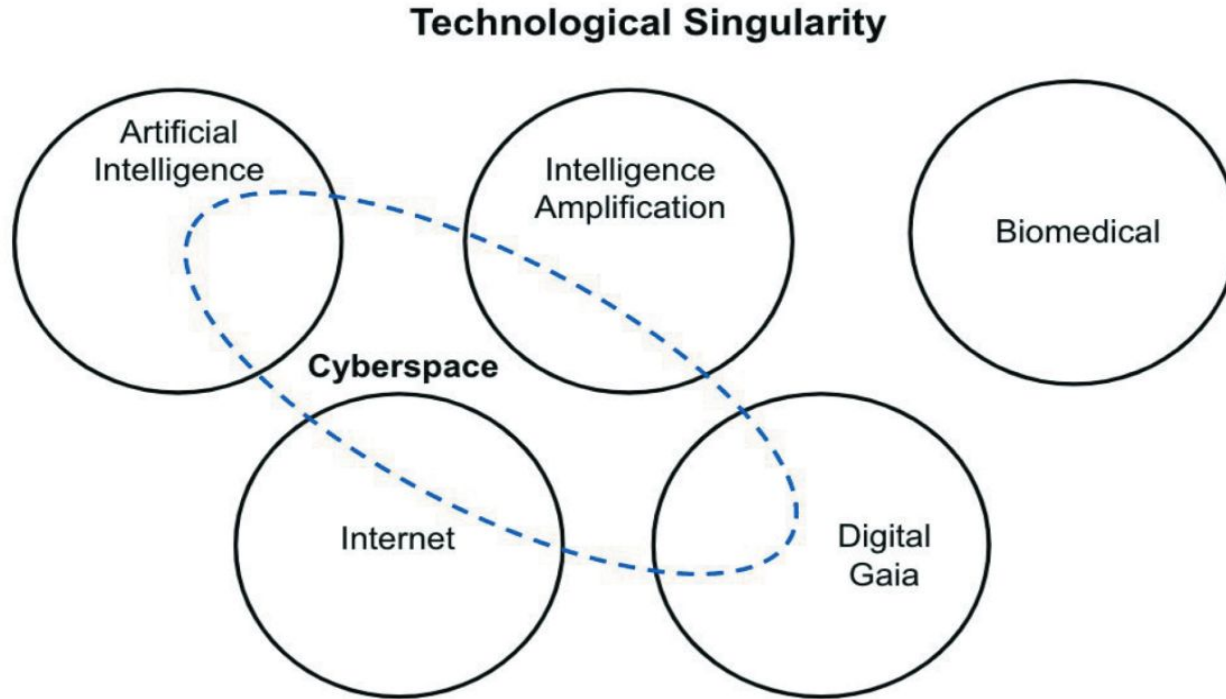
Vinge's Singularity Hypothesis

In *The Coming of Technological Singularity, 1993*, and *The Signs of Singularity, IEEE Spectrum*, Vinge predicted the possible paths for singularity:

- The AI Scenario: We create superhuman artificial intelligence (AI) in computers.
- The IA Scenario: We enhance human intelligence through human-to-computer interfaces—that is, we achieve intelligence amplification (IA).
- The Biomedical Scenario: We directly increase our intelligence by improving the neurological operation of our brains.
- The Internet Scenario: Humanity, its networks, computers, and databases become sufficiently effective to be considered a superhuman being.
- The Digital Gaia Scenario: The network of embedded microprocessors becomes sufficiently effective to be considered a superhuman being.

Predicted the oncoming of Technological Singularity in **2005-2030**

Proposed Research Work



Proposed Research Work

Computers, Materials & Continua
DOI:10.32604/cmc.2021.015250
Article

Tech Science Press

Analyzing Some Elements of Technological Singularity Using Regression Methods

Ishaani Priyadarshini^{1,*}, Pinaki Ranjan Mohanty² and Chase Cotton¹

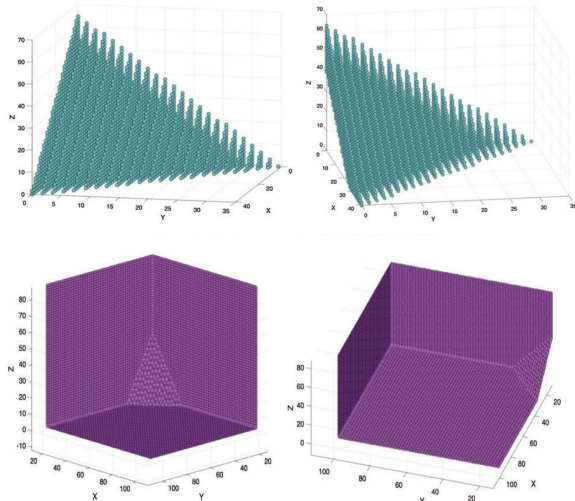


Figure 9: Solution Space for Eq. (1) (2045)

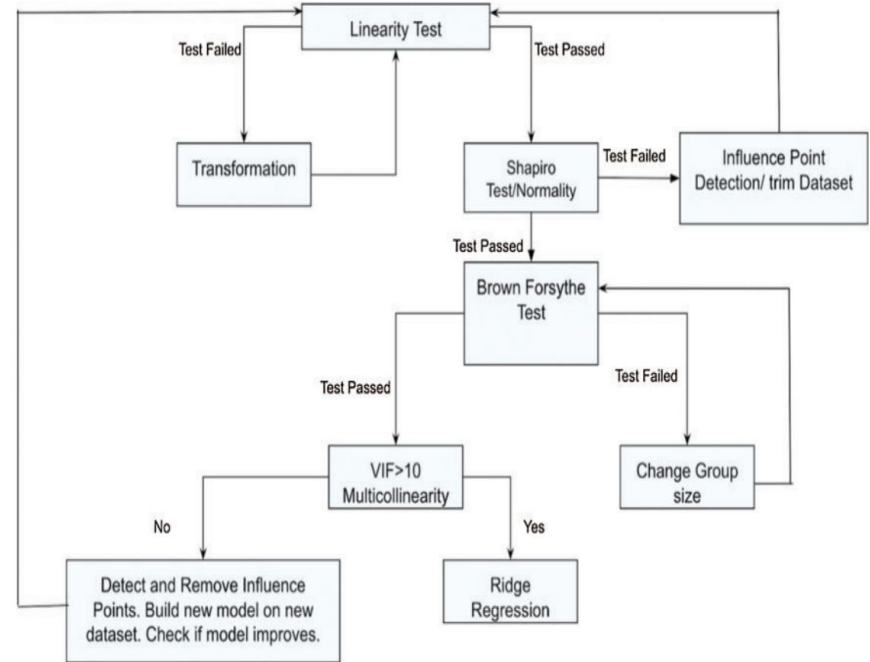


Figure 1: Steps for the analysis conducted

Experiments with OCR and Facial Emotions

