



UNIVERSITY OF ZAGREB

Faculty of Electrical
Engineering and
Computing



FunTomP
FunctionalizedTomatoProducts

Machine learning and spectroscopy

Sanda Pleslić, Eda Jovičić, Franko Katalenić

University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000
Zagreb, Croatia

Nadica Maltar-Strmečki, Kristina Smokrović

Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

FunTomP Project Meeting Valencia 9-10 November 2023



FunTomP



This project has received funding
from the European Union's PRIMA
programme under grant agreement
No 2032





<https://itsoftware.com.co/content/structured-vs-unstructured-data/>

Data – Information - Knowledge

John Naisbitt (1982) wrote: **“We are drowning in information but starved for knowledge”.**

The data is raw, it simply exists and has no other meaning and meaning than its existence.

Information is data that is given meaning through a relational connection.

Knowledge is a suitable collection of information and such that it can be considered useful.

Knowledge are patterns in data or models which explain them.



Machine learning

<https://www.udacity.com/course/machine-learning--ud262>

Machine learning must be one of the fastest growing fields in computer science.

It is not only that the data is continuously getting “bigger,” but also the theory to process it and turn it into knowledge.

In various fields of science, from astronomy, physics, chemistry to biology, but also in everyday life, as digital technology increasingly infiltrates our daily existence, as our digital footprint deepens, more data is continuously generated and collected.



Machine learning

<https://www.udacity.com/course/machine-learning--ud262>

Whether scientific or personal, data that just lies dormant passively is not of any use, and smart people have been finding ever new ways to make use of that data and turn it into a useful product or service.

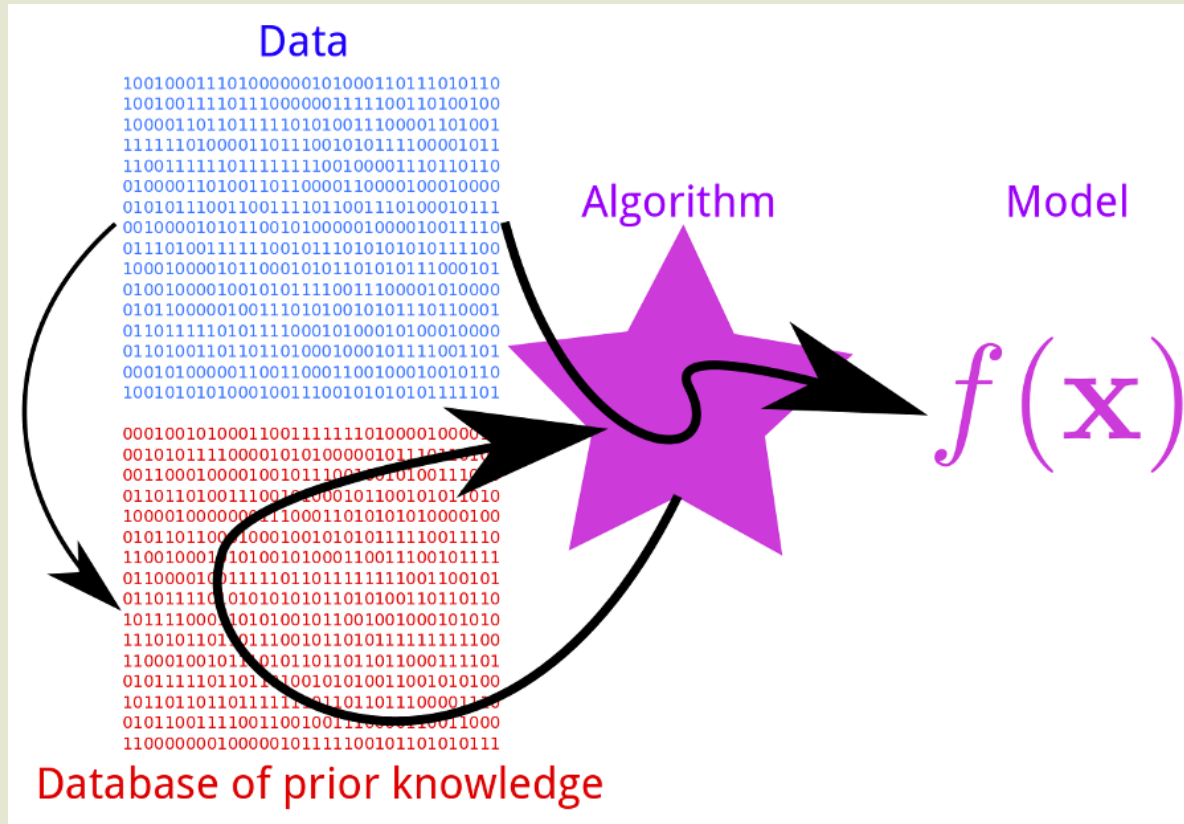
In this transformation, **machine learning** plays a larger and larger role.

Machine learning

The **machine learning process** begins with **observations of data**, such as examples, direct experience or instruction.

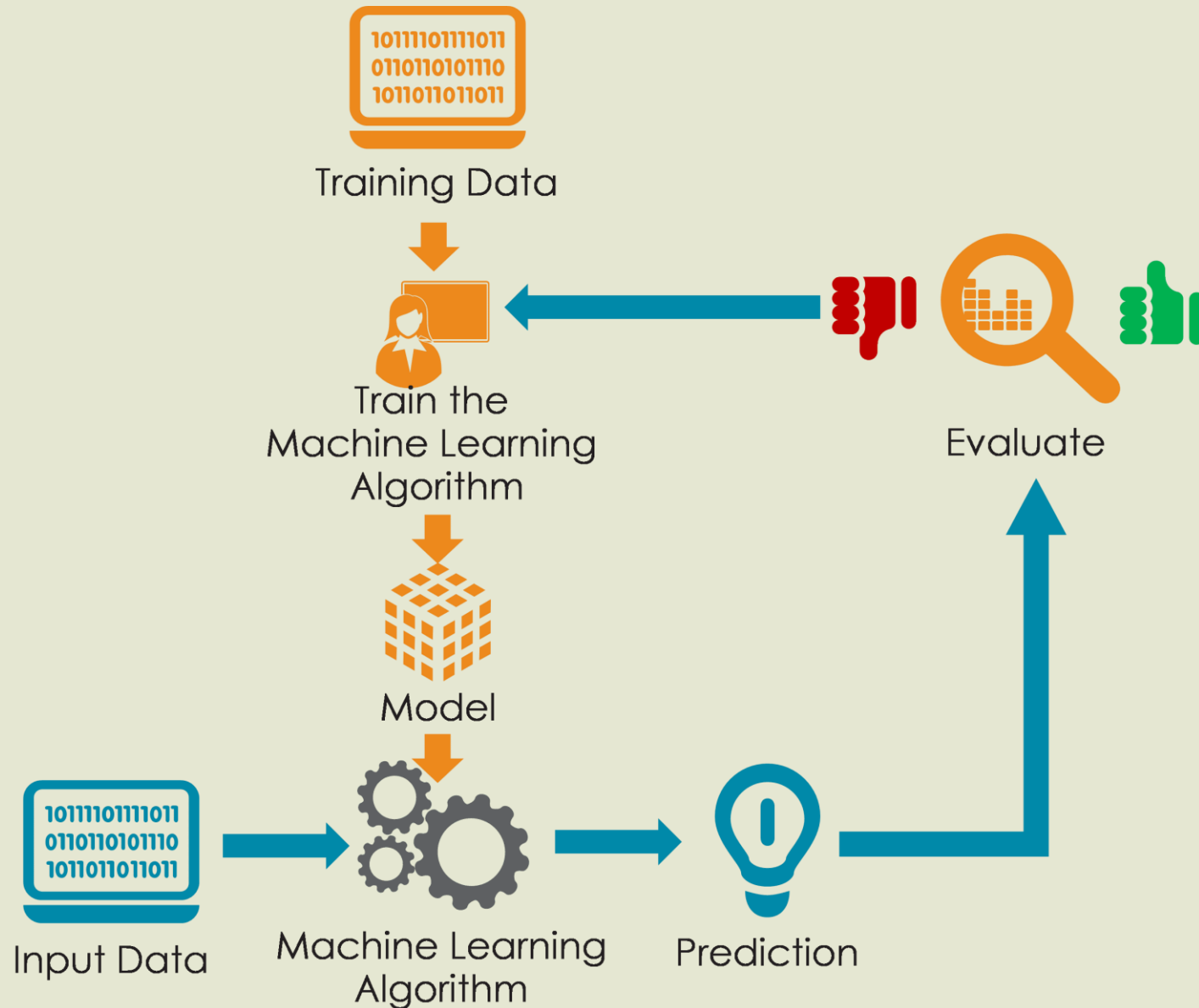
ML looks for **rules, patterns or models in data** so it can later make inferences based on the examples provided.

The primary aim of ML is **to allow computers to learn autonomously without human intervention or assistance and adjust actions accordingly.**



<https://medium.com/@kaykwaalk/difference-between-an-algorithm-and-a-model-in-machine-learning-5d6b7297cce9>

Machine learning



ML has proven valuable because it can solve problems at a speed and scale that cannot be duplicated by the human mind alone.

With massive amounts of computational ability behind a single task or multiple specific tasks, machines can be trained to identify **patterns** in and **relationships** between input data and automate routine processes.

Machine learning: Data is key



<https://www.businessprocessincubator.com/content/safeguarding-data-ensuring-privacy/>

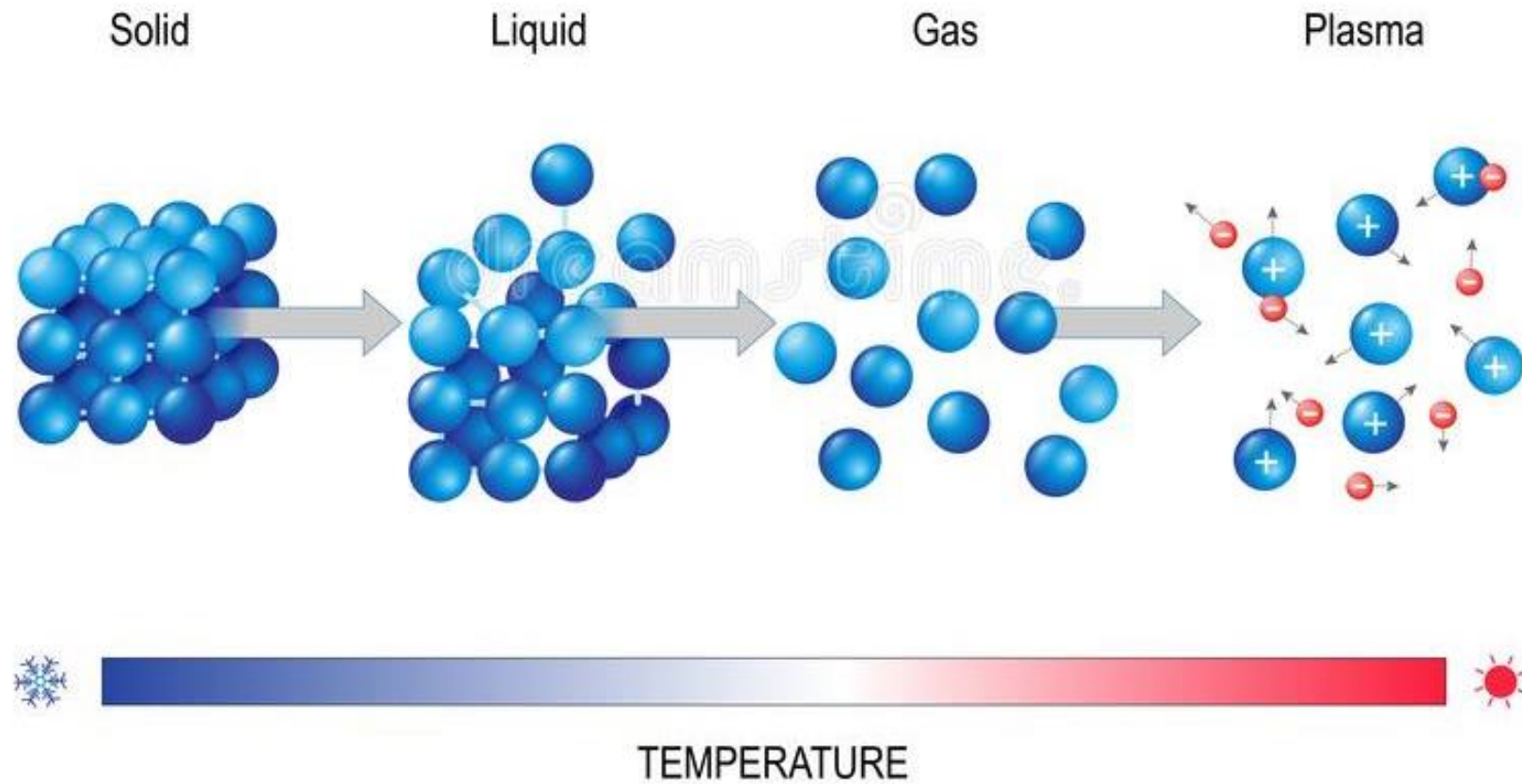
The **algorithms** that drive machine learning are critical to success.

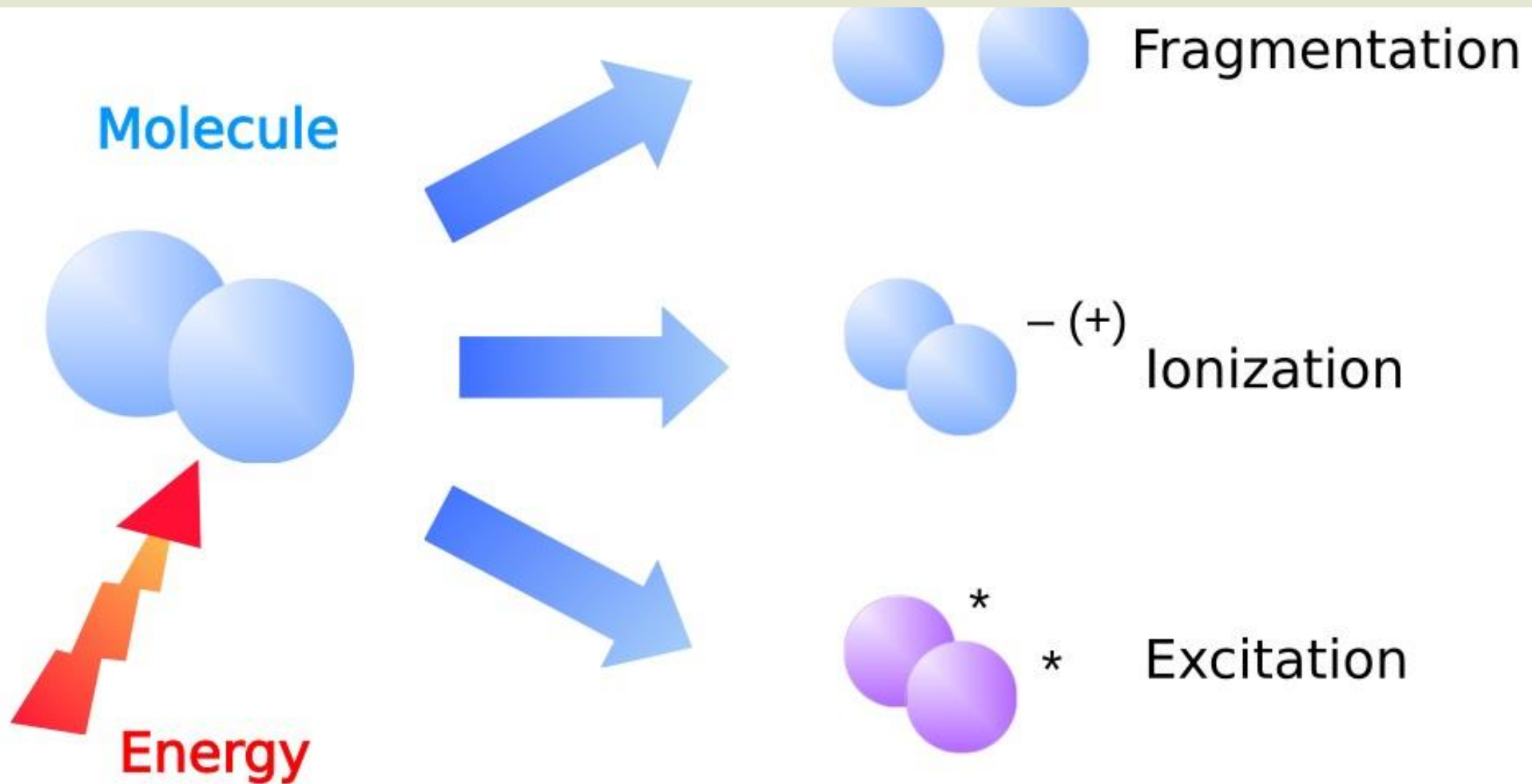
ML algorithms build a mathematical model-based sample data, known as **“training data”** to make **predictions or decisions** without being explicitly programmed to do so.

This can reveal trends within data that information businesses can use **to improve decision making, optimize efficiency and capture actionable data at scale.**

The goal of ML: to build models that generalize well.

STATES OF MATTER





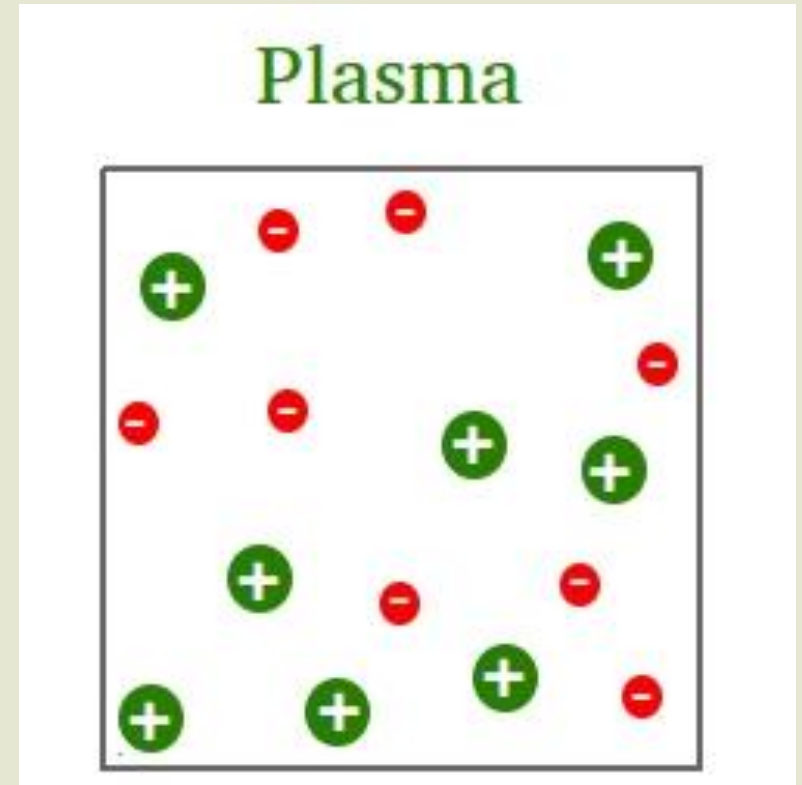
Plasma – state of matter

Plasma, in physics, an electrically conducting medium in which there are roughly equal numbers of positively and negatively charged particles, produced when the atoms in a gas become ionized.

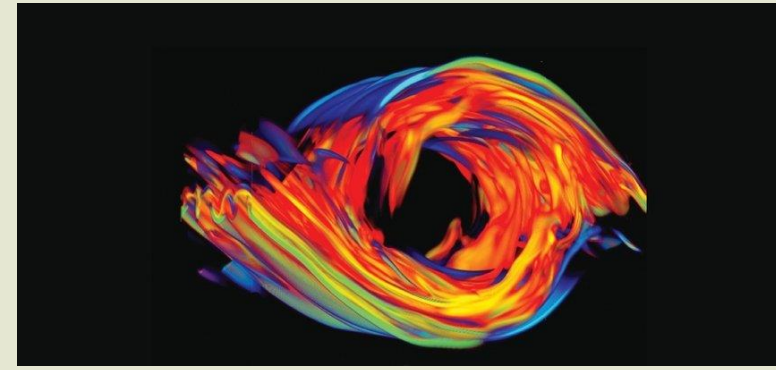
It is sometimes referred to as **the fourth state of matter**, distinct from the solid, liquid and gas states.

The negative charge is usually carried by **electrons**.

The positive charge is typically carried by atoms or molecules that are missing those same electrons, they are **ions**.



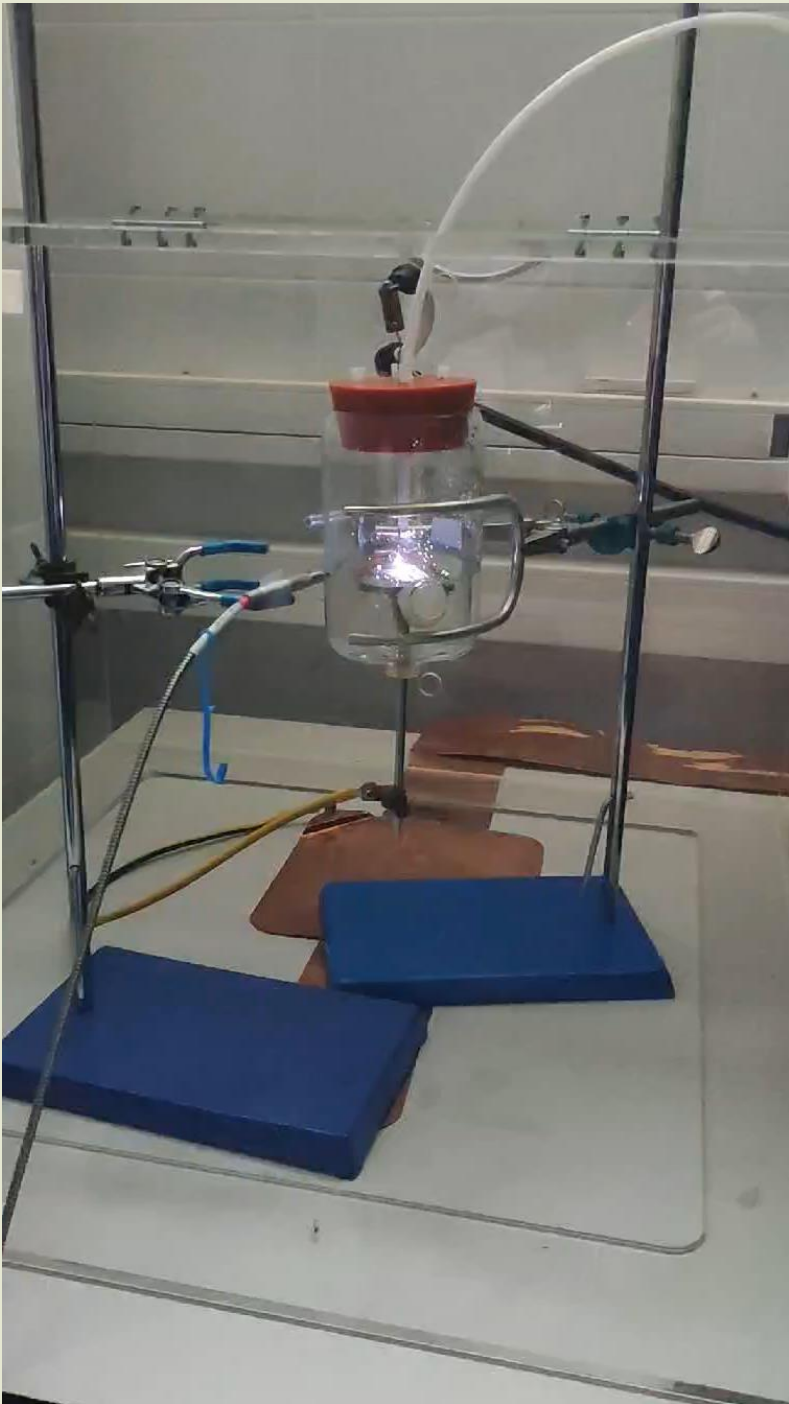
Plasma – state of matter



In some cases, **electrons** missing from one type of atom or molecule become attached to another component, resulting in a plasma containing both positive and negative **ions**.

The uniqueness of the plasma state is due to the importance of **electric and magnetic forces** that act on a plasma in addition to such forces as gravity that affect all forms of matter.

Since these electromagnetic forces can act at large distances, a plasma will act collectively much like a **fluid** even when the particles seldom collide with one another (magnetohydrodynamical description of **plasma fluid** - MHD).



The **experimental set-up** consisted of a glass reactor (300-1000 mL) with a point-to-plate electrode configuration.

A medical stainless-steel needle was used as a HV electrode and stainless-steel electrode was a ground electrode.

Plasma was generated in liquid and gas phase through the capacitor charged with the pulsed HV power supply (IMPEL HVG60/1) to create electrical discharges up to maximal 50 kV; frequency range was 0-300 Hz.

Plasma diagnostics was done with Ocean FX extended range model (OCEAN-FX-XR1-ES) of **optical emission spectrometer** (OES).

The emission of a certain volume in the plasma chamber was delivered by an optical fiber to the spectrometer.

Optical emission spectrometer (OES)

FunTomP

FunctionalizedTomatoProducts



FunTomP



[OCEAN-FX-XR1-ES](#)

Ocean FX extended-range model spectrometer preconfigured for 200-1025 nm wavelength range, with 25 Åμm slit and lens for enhanced sensitivity

[P600-2-SR74-UV](#)

600 μm Fiber, solarization-resistant, 2 m UV/VIS Collimating Lens, 200-2500 nm

[OCEANVIEW](#)

OceanView spectroscopy software with graphical user interface; accessible by download only from our secure server

OPTICAL EMISSION SPECTROSCOPY (OES) FOR PLASMA ANALYSIS

Plasma composition is a critical parameter in all types of plasma applications.

“Optical emission spectroscopy” is a common term for techniques that measure **optical absorption/emission spectra**, i.e., the light being emitted when a sample is being excited.

OES is an important technique in plasma analysis, where it enables **accurate non-contact characterization** of the color of the plasma and thus the species present.

Crucially, **OES is compatible with thermal and non-thermal plasma environments and applications** that are sensitive to contamination because it is the ideal **non-invasive measurement technique for plasma diagnostics** in hostile plasma environments while **preventing contamination**.

Additionally, OES provides **near-real-time feedback**.

OPTICAL EMISSION SPECTROSCOPY (OES) FOR PLASMA ANALYSIS

However, plasmas are typically characterized by a **large number of closely-spaced emission peaks**.

Separating these peaks requires an OES system with **high spectral resolution**.

Previous research on OES measurements of plasma processes has largely focused on plasma monitoring, species identification, and determination of the **electron temperature T_e and electron density n_e** .

In order to use the OES method to determine these parameters, one usually applies the so-called **line intensity ratio method**, which requires a **relative intensity calibrate spectroscopic system** and a **suitable physics-based model for excited species in the plasmas to be investigated**.

Line Intensity Ratio

In optical plasma diagnostics, line intensity ratio method is one of method to determine electron temperature and density of plasma. The intensity of radiation, I_{ki} , is expressed by

$$I_{ki} = E(k, i)_{ki} N_k A_{ki}$$

E_{ki} : Difference between energy levels E_k and E_i ($E_k > E_i$)

A_{ki} : Transition probability

N_k : Population of upper energy state of k .

Line intensity ratio of plasma is expressed by

$$\mathbb{L} = \frac{I_{ki}}{I_{pq}} = \frac{\lambda_{pq} N_k A_{ki}}{\lambda_{ki} N_p A_{pq}}$$

λ_{ki} : Wavelength corresponding to the transition from energy state k to energy state i .

In experiment, the optical system cannot collect whole radiation from plasma volume, the transmittance of optical system has a dependency on wavelength, and quantum efficiency of detecting system is also dependent on a wavelength. Thus, the measured line intensity ratio needs to be relatively calibrated for whole optical system and detector.

Boltzmann plot

The spectroscopic data also was used for estimating the temperature of the produced plasma and it is carried out using **Boltzmann plot**:

$$\ln \left(\frac{I}{\lambda g A} \right) = - \frac{1}{k_B T} E + \ln \left(\frac{4\pi Z}{h c n_0} \right)$$

I is intensity of the spectral line,

λ is wavelength of spectral line,

g is statistical weight of the energy level,

A is transition probability,

k_B is Boltzmann constant,

T is temperature,

E is energy level of the upper state for emission,

Z is partition function,

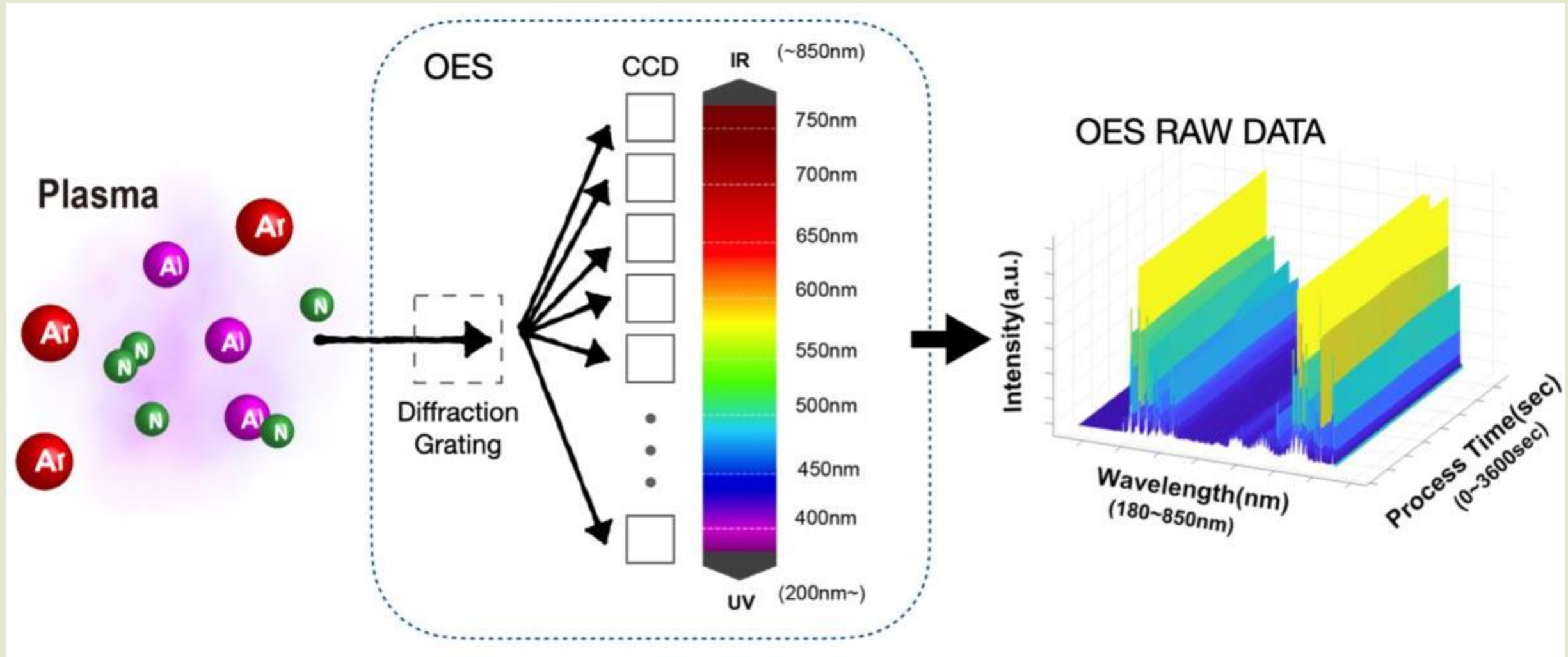
n_0 is total species population,

k_B is Boltzmann constant,

h is Planck constant and

c is speed of light.

Spectroscopy



OES spectra





NIST

National Institute of Standards and Technology

Physical Meas. Laboratory

<https://www.nist.gov/>

is used for identification and confirmation of the spectral lines.

NIST: Atomic Spectra Database | x +

← → ↻ 🔒 physics.nist.gov/PhysRefData/ASD/lines_form.html

Google 🔗 ☆ □ Pauzirano ⋮



DATA
LINES LEVELS

INFORMATION
List of SPECTRA
GROUND STATES & IONIZATION ENERGIES
Bibliography Help



National Institute of Standards and Technology
Physical Meas. Laboratory

Main Parameters

Spectrum

e.g., Fe I or Na;Mg; Al or mg i-iii or 198Hg I

Limits for Wavelengths ▾ Lower:

Upper:

Wavelength Units: nm ▾

Reset input Retrieve Data Show Graphical Options Show Advanced Settings

Can you please provide some [feedback](#) to improve our database?



DATA

LINES

LEVELS

INFORMATION

List of
SPECTRAGROUND STATES &
IONIZATION ENERGIES

Bibliography

Help

NIST Atomic Spectra Database Lines Data

Ar (all spectra): 1300 Lines of Data Found

Example of how to reference these results:

Kramida, A., Ralchenko, Yu., Reader, J., and NIST ASD Team (2022). *NIST Atomic Spectra Database* (ver. 5.10), [Online]. Available: <https://physics.nist.gov/asd> [2023, November 7]. National Institute of Standards and Technology, Gaithersburg, MD. DOI: <https://doi.org/10.18434/T4W30F>

[BibTex Citation](#) (new window)

Wavelength range: 350 - 850 nm

Wavelength in: vacuum below 200 nm, air between 200 and 2000 nm, vacuum above 2000 nm

Highest relative intensity: 35000

Query NIST Bibliographic Databases for
Ar (new window)

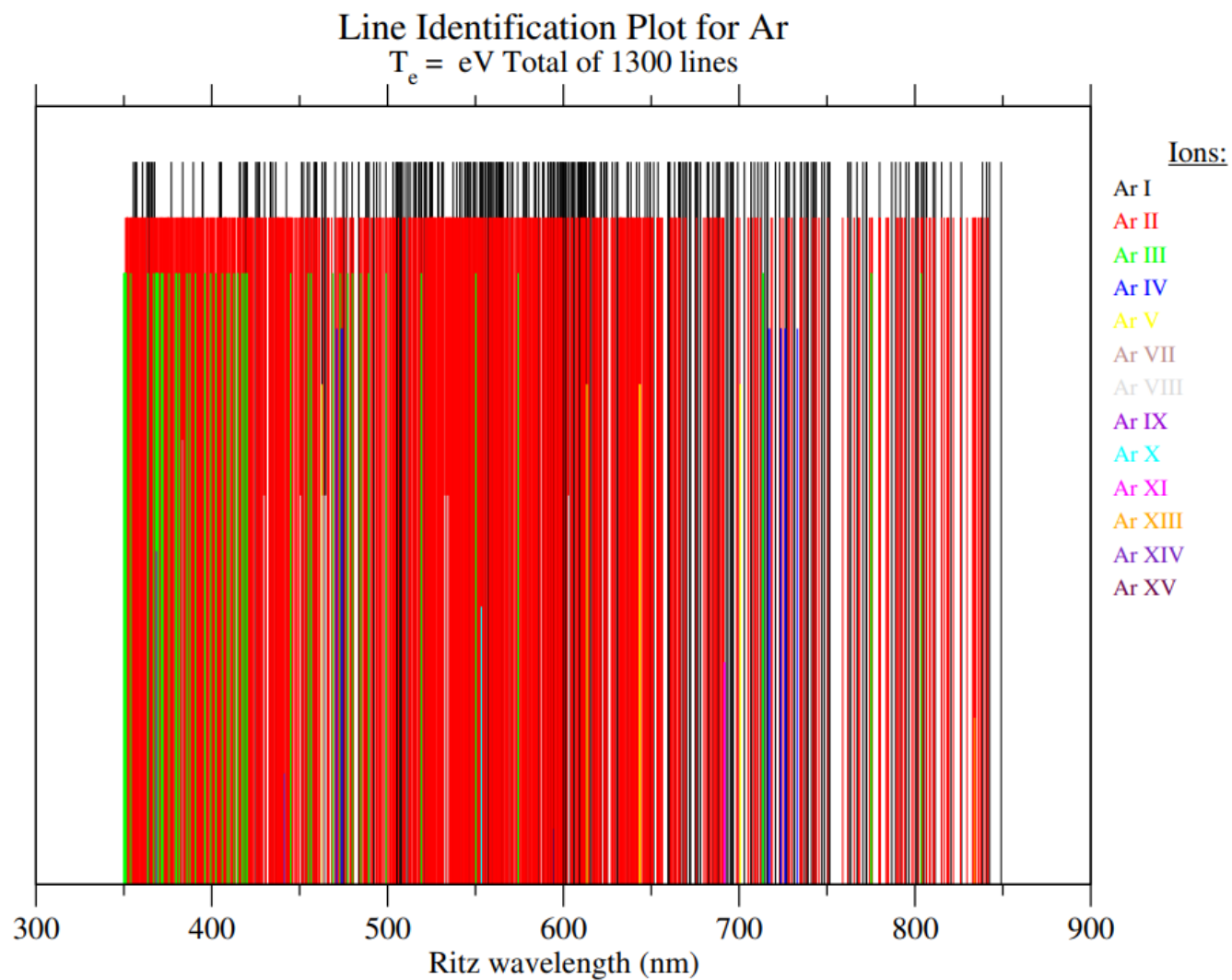
[Ar Energy Levels](#)[Ar Line Wavelengths and Classification](#)[Ar Transition Probabilities](#)

Ion	Observed Wavelength Air (nm)	Unc. (nm)	Ritz Wavelength Air (nm)	Unc. (nm)	Rel. Int. (?)	A_{ki} (s ⁻¹)	Acc.	E_i (cm ⁻¹)	E_k (cm ⁻¹)	Lower Level Conf., Term, J	Upper Level Conf., Term.
Ar III	350.058	0.005	350.057415	0.000017	5	2.6e+07	D	196 590.6352 -	225 149.2072	3s ² 3p ³ (² D°)4s ³ D° 1	3s ² 3p ³ (² D°)4p ³ D°
Ar III	350.26828	0.00001	350.268291	0.000016	93	4.3e+07	D	196 615.2125 -	225 156.5915	3s ² 3p ³ (² D°)4s ³ D° 2	3s ² 3p ³ (² D°)4p ³ D°

Range 350-850 nm

NIST ASD Output: Lines															
/home/www/htdocs/PhysRefDat															
physics.nist.gov/cgi-bin/ASD/lines1.pl?spectra=Ar&limits_type=0&low_w=350&upp_w=850&unit=1&submit=Retrieve+Data&de=0...															
Ion	Observed Wavelength Air (nm)	Unc. (nm)	Ritz Wavelength Air (nm)	Unc. (nm)	Rel. Int. (?)	A_{ki} (s ⁻¹)	Acc.	E_i (cm ⁻¹)	E_k (cm ⁻¹)	Lower Level Conf., Term, J			Upper Level Conf., Term		
Ar III	350.058	0.005	350.057415	0.000017	5	2.6e+07	D	196 590.6352 -	225 149.2072	3s ² 3p ³ (² D°)4s	³ D°	1	3s ² 3p ³ (² D°)4p	³ D	
Ar III	350.26828	0.00001	350.268291	0.000016	93	4.3e+07	D	196 615.2125 -	225 156.5915	3s ² 3p ³ (² D°)4s	³ D°	2	3s ² 3p ³ (² D°)4p	³ D	
Ar III	350.35894	0.00001	350.358940	0.000014	468	1.2e+08	D	196 615.2125 -	225 149.2072	3s ² 3p ³ (² D°)4s	³ D°	2	3s ² 3p ³ (² D°)4p	³ D	
Ar III	350.93333	0.00001	350.933330	0.000017	245			180 678.3147 -	209 165.6077	3s ² 3p ³ (⁴ S°)4s	³ S°	1	3s ² 3p ³ (⁴ S°)4p	³ P	
Ar II	350.97779	0.00001	350.977794	0.000009	447	2.55e+08	B	155 708.1075 -	184 191.7917	3s ² 3p ⁴ (³ P)4p	⁴ P°	1/2	3s ² 3p ⁴ (³ P)4d	⁴ D	
Ar III	351.11485	0.00001	351.114850	0.000017	1175			180 678.3147 -	209 150.8807	3s ² 3p ³ (⁴ S°)4s	³ S°	1	3s ² 3p ³ (⁴ S°)4p	³ P	
Ar III	351.16671	0.00001	351.166713	0.000014	89	2.6e+07	D	196 680.8461 -	225 149.2072	3s ² 3p ³ (² D°)4s	³ D°	3	3s ² 3p ³ (² D°)4p	³ D	
Ar III	351.42004	0.00001	351.420041	0.000017	724			180 678.3147 -	209 126.1544	3s ² 3p ³ (⁴ S°)4s	³ S°	1	3s ² 3p ³ (⁴ S°)4p	³ P	
Ar II	351.43874	0.00001	351.438733	0.000005	490	1.36e+08	B	155 351.1206 -	183 797.4473	3s ² 3p ⁴ (³ P)4p	⁴ P°	3/2	3s ² 3p ⁴ (³ P)4d	⁴ D	
Ar II	351.78897	0.00010	351.788887	0.000007	4	7.e+04	E	132 630.7277 -	161 048.7411	3s ² 3p ⁴ (³ P)3d	⁴ D	3/2	3s ² 3p ⁴ (³ P)4p	⁴ S	
Ar II	351.893	0.004	351.8922	0.0014	1w,bl			190 592.2306 -	219 001.90	3s ² 3p ⁴ (³ P)4d	² P	3/2	3s ² 3p ⁴ (¹ D ₂)5f	² [	
Ar II	351.99933	0.00001	351.999331	0.000004	191	5.2e+07	B	157 673.4134 -	186 074.4375	3s ² 3p ⁴ (³ P)4p	⁴ D°	5/2	3s ² 3p ⁴ (³ P)4d	⁴ F	
Ar II	352.12600	0.00001	352.125990	0.000005	100	2.27e+07	B	157 234.0196 -	185 624.8282	3s ² 3p ⁴ (³ P)4p	⁴ D°	7/2	3s ² 3p ⁴ (³ P)4d	⁴ F	

Range 350-850 nm

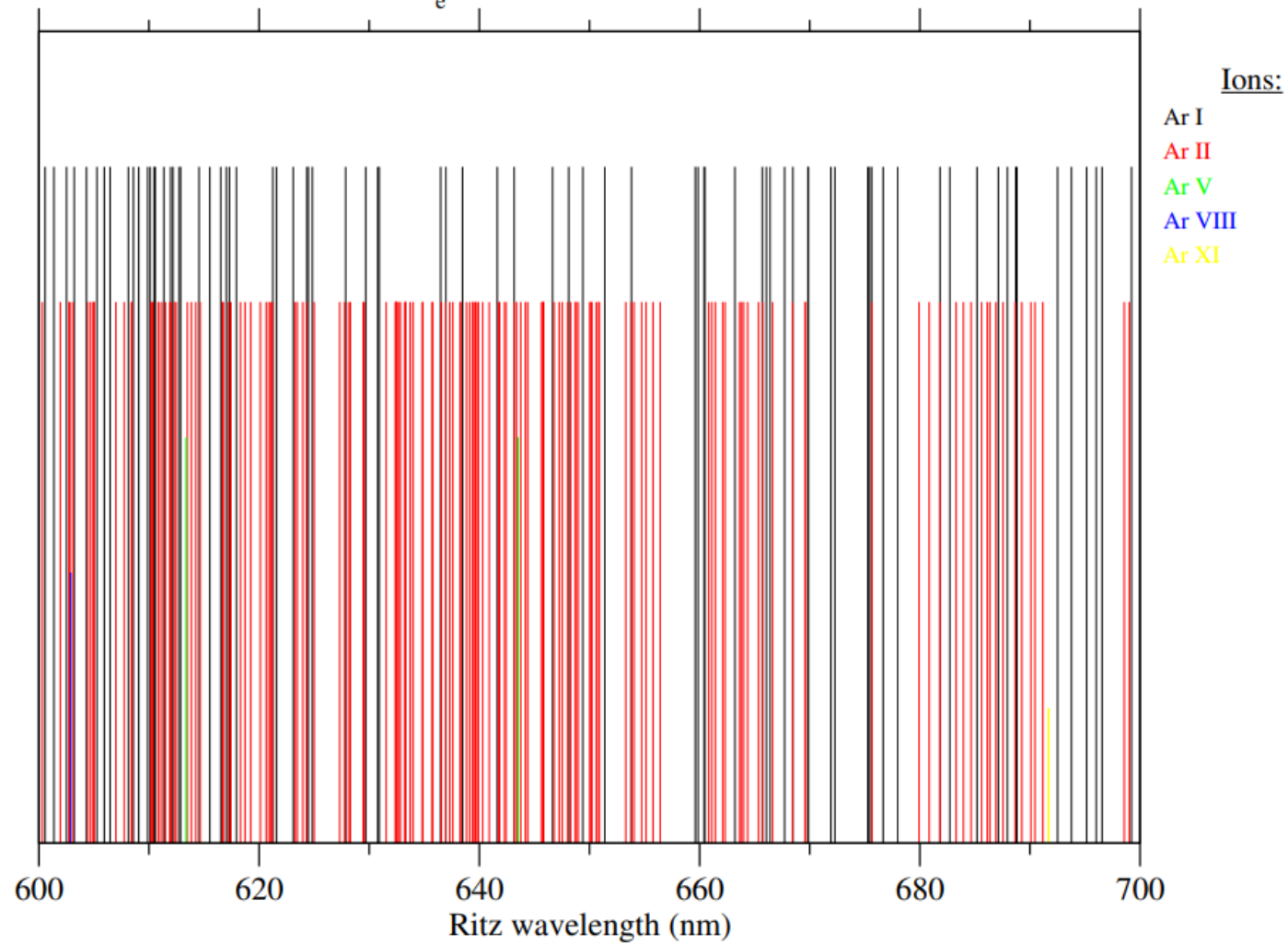


Range 600-700 nm

NIST ASD Output: Lines													
physics.nist.gov/cgi-bin/ASD/lines.pl?spectra=Ar&limits_type=0&low_w=600&upp_w=700&unit=1&submit=Retrieve+Data&de=0<e_...													
Ion	Observed Wavelength Air (nm)	Unc. (nm)	Ritz Wavelength Air (nm)	Unc. (nm)	Rel. Int. (?)	A_{ki} (s ⁻¹)	Acc.	E_i (cm ⁻¹)	E_k (cm ⁻¹)	Lower Level Conf., Term, J			Upper Level Conf., Term, J
Ar II	600.3170	0.0008	600.31857	0.00006	1			158 167.7999 - 174 821.0098		3s ² 3p ⁴ (³ P)4p	⁴ D°	³ / ₂	3s ² 3p ⁴ (¹ D)3d ² P
Ar I			600.5725	0.0010		1.4e+05	D+	107 289.7001 - 123 935.870		3s ² 3p ⁵ (² P° _{1/2})4p	² [³ / ₂]	2	3s ² 3p ⁵ (² P° _{3/2})8s ² [³ / ₂]°
Ar I			601.3677	0.0010		1.4e+05	D+	105 462.7596 - 122 086.9164		3s ² 3p ⁵ (² P° _{3/2})4p	² [⁵ / ₂]	3	3s ² 3p ⁵ (² P° _{3/2})5d ² [³ / ₂]°
Ar II	601.9493	0.0008	601.94965	0.00005	4			172 829.6539 - 189 437.7396		3s ² 3p ⁴ (¹ D)3d	² D	³ / ₂	3s ² 3p ⁴ (³ P)5p ⁴ P°
Ar I	602.5150	0.0010	602.5150	0.0010	5	9.0e+05	D+	107 289.7001 - 123 882.203		3s ² 3p ⁵ (² P° _{1/2})4p	² [³ / ₂]	2	3s ² 3p ⁵ (² P° _{1/2})7s ² [¹ / ₂]°
Ar II	602.72513	0.00007	602.72517	0.00003	7			173 347.9158 - 189 934.6322		3s ² 3p ⁴ (¹ D)4p	² D°	³ / ₂	3s ² 3p ⁴ (³ P)4d ² P
Ar II	602.8220	0.0008	602.823273	0.000015	1			190 511.2674 - 207 095.2846		3s ² 3p ⁴ (³ P)5p	² D°	⁵ / ₂	3s ² 3p ⁴ (³ P)7s ⁴ P
Ar VIII	602.98	0.08	602.9	0.4				1 070 277. - 1 086 860.		2p ⁶ 9f	² F°	⁷ / ₂	2p ⁶ 10g ² G
Ar II	603.0844	0.0008	603.08278	0.00004	11			191 169.5946 - 207 746.4757		3s ² 3p ⁴ (³ P)5p	⁴ S°	³ / ₂	3s ² 3p ⁴ (³ P)6d ⁴ D
Ar I	603.2127	0.0010	603.2127	0.0010	70	2.46e+06	C	105 462.7596 - 122 036.0704		3s ² 3p ⁵ (² P° _{3/2})4p	² [⁵ / ₂]	3	3s ² 3p ⁵ (² P° _{3/2})5d ² [⁷ / ₂]°
Ar I	604.3223	0.0010	604.3223	0.0010	35	1.47e+06	C	105 617.2700 - 122 160.1502		3s ² 3p ⁵ (² P° _{3/2})4p	² [⁵ / ₂]	2	3s ² 3p ⁵ (² P° _{3/2})5d ² [⁷ / ₂]°
Ar II	604.4468	0.0008	604.44783	0.00009	7			192 712.058 - 209 251.5031		3s ² 3p ⁴ (³ P)4d	² D	³ / ₂	3s ² 3p ⁴ (¹ D ₂)4f ² [3]°
Ar II	604.68977	0.00007	604.68967	0.00005	10			174 409.8916 - 190 942.7220		3s ² 3p ⁴ (¹ D)3d	² P	³ / ₂	3s ² 3p ⁴ (³ P)5p ² S°

Range 600-700 nm

Line Identification Plot for Ar
 $T_e =$ eV Total of 236 lines



Range 610-620 nm

NIST NIST ASD Output: Lines

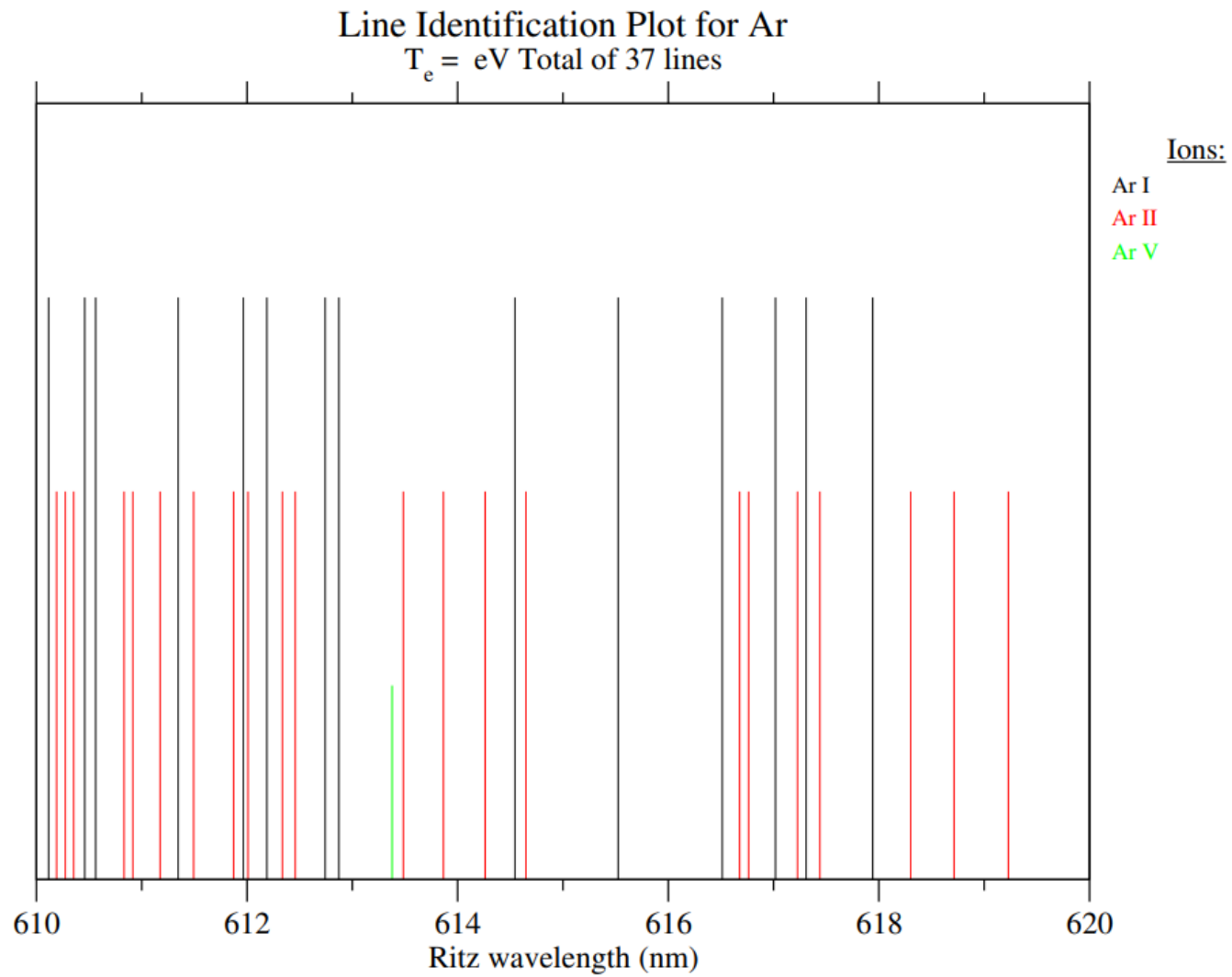
/home/www/htdocs/PhysRefDat

+

← → ↺

physics.nist.gov/cgi-bin/ASD/lines1.pl?spectra=Ar&limits_type=0&low_w=610&upp_w=620&unit=1&submit=Retrieve+Data&de=0...

Range 610-620 nm



Role of machine learning for optical emission spectroscopy

The emphasis of mentioned models is the **identification of major production and depopulation processes under different plasma discharge conditions and for different kinds of excited species.**

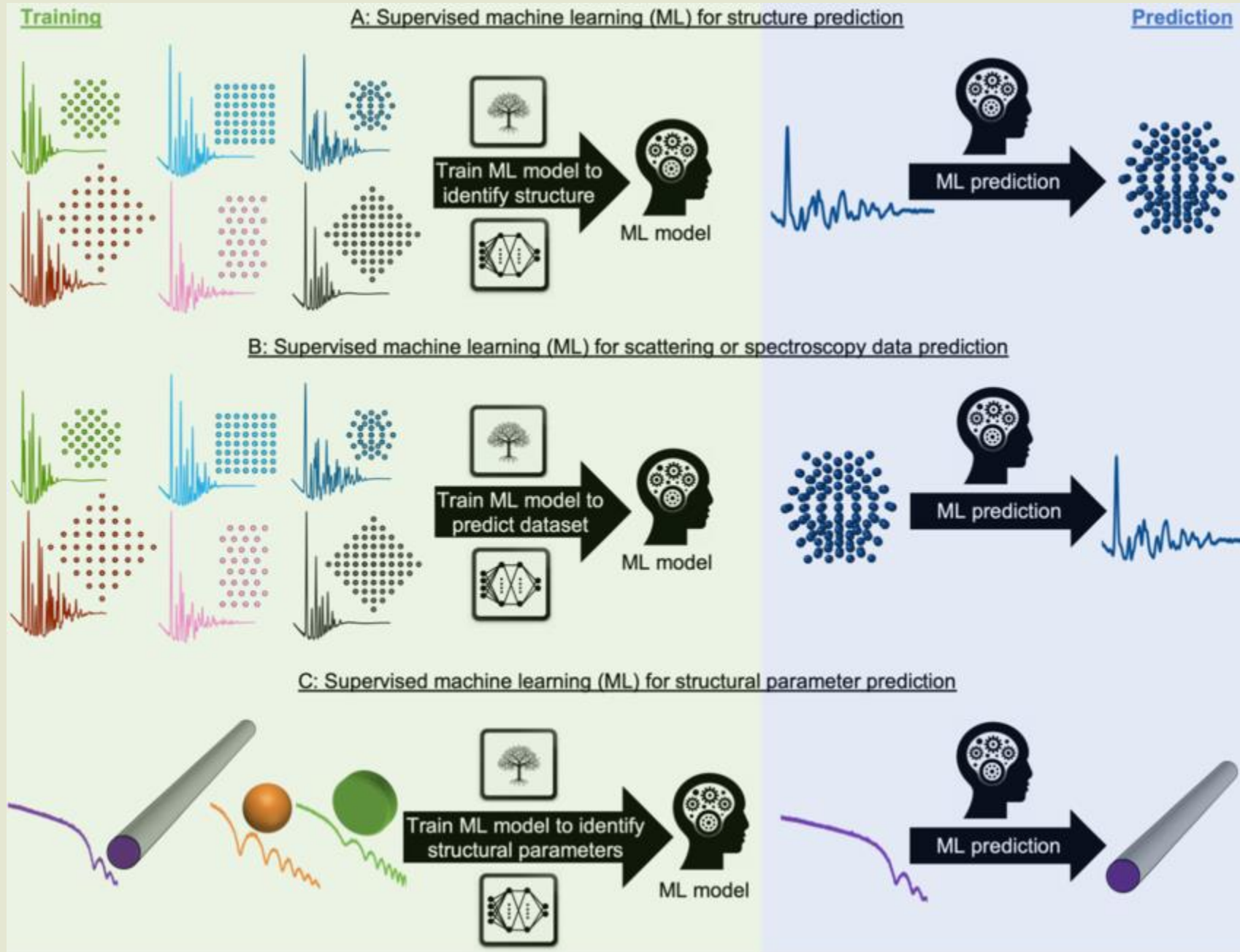
The quality of the modelling results, which determines the accuracy of diagnostic results by the line intensity ratio method, depends on the existence and quality of the data used.

The line intensity ratio method is expected to be further developed in the future.

Many studies have demonstrated the **effectiveness of ML for OES.**

ML for OES

The fundamental idea in ML is that, for many applications, **training a computer algorithm for predicting or finding patterns** in the behavior of a complex system by **observing many input–output samples** of its behavior can be significantly simpler than developing physics-based models.



FunTomP

FunctionalizedTomatoProducts



FunTomP

**Thank you for your
attention!**



UNIVERSITY OF ZAGREB



Faculty of Electrical
Engineering and
Computing



PRIMA
PARTNERSHIP FOR RESEARCH AND INNOVATION
IN THE MEDITERRANEAN AREA



This project has received
funding from the
European Union's PRIMA
programme under grant
agreement No 2032