

Predicting Crime in Athens: A Machine Learning Approach

<u>Speaker :</u> Maria Eleni Vardaki

Associate Professor: PhD Constantinos Mourlas, PhD Catherine Sotirakou





• Big Data & Artificial Intelligence (AI)



A model to predict crime

Presentation modules

- Crime in Media
- Data Mining & Big Data
- Data Journalism



Presentation modules

- Machine Learning
- Algorithms
- Our model from $A \rightarrow Z$



Crime: Offence punishable by law

- Place
- Time
- Social-economic solutions



"News": Things that happen during the day



Social Media



Source: Crowdynews



Framing theory:

- Journalists make decisions
- Manipulate audiences and feelings
- Visual & Verbal framing are highly important



Greek journalists:

• Create & cultivate stereotypical perceptions

• Intensify xenophobia

Data mining



Source: <u>Noteworthy — The Journal Blog</u>





- High volume of structured or unstructured data
- <u>3Vs:</u>
- ✓ Volume
- ✓ Velocity
- ✓ Variety





Data Journalism

"Drawning conclusions through the management of digital data with the aim of their final capture through an Infographic type chart" Veglis & Bratsas



Source: Wikimedia



Data Journalism

								(m		
Area second and second and a share second in	DAY SCHOOLS Falablahara	Total I	Gipla.	Tels!	Kap	Remarks.	Metters, Temisman and Report, Through		BLAGRAM OF THE CAUSES OF MORTALI	ITY
beatly affected in the back of the band, that for me				-	-	the second s	slines .			1.
lef he had remainer to blooding, and a did in a	Gransser School	\$99		144	1900	1	LEE Jonathon, of Sunderland, in the canal	APRIL 1855 TO MARCH 5856	IN THE ARMY IN THE EAST.	APRIL 1854 TO MARCH 1855
carden had her arm andhods lifted on he for offer	Blue Cost dates	885.	** **	80	2000	Faught, clothed and baarded.	Darham, grocos, d. c.; May 12, 19, Jane	NOTE that is in the line of th		THE REAL PROPERTY AND AND A
and first in becoming for some up by its entert,	Green Coat ditto	- 59		- 50	200	Tanght and clotherd.	12, at Gaddhall, Loudon, Atta Moura O	1986		
coloristican of his Theff and cloud they. The	Collegiste Church dater	****	-50	-50	-40	And offertory manay i da, da. 17	Huddon and tistly, Thransoton street.	100/00/201		and the second s
whilst on the best but any much struct on the arm	Straugeways ditte	101	****	30	100		PAVN Thomas, and John Daniel Pays, of C	ALC: NO REPORT OF		
war fallenter also are merry stamed. A lady						(Suppose)-Taucht and elotherd.	ton-street, in the city of London, watchoose	100000000000000000000000000000000000000		2012
the firsters in contact with the time, 5/8 in-	St. Mary's ditto.	12	12	24	40	Funds arising from Sacramontal	d. c.; Mar 5, Pf. Jane 9, at 1, at Calls	and the second second		ALC: ALC: ALC: ALC: ALC: ALC: ALC: ALC:
consider the contact with the advice a sequences			1.1.1.1			Officiant.	Att. Mr. Elizaberte, Barinchall street	ALTER AND AND ADDRESS AND		ACCOUNT OF ANY OF A DECK
control will be a sagest cleaterst spark. The	St. John's dista.			. 97		(Nepport)- Expenses raised by su-	SMITH John non or lots of Patheesters in	and the second second second second second		JUHE 3
View View valered the house of Mrs. Clare, in	St. Paul's ditta	29		-90(1		Huntary Salaryintian.	demanus, in the county of York stores it	-STREET, CONTRACTOR STREET, ST		Contraction in the second second second
rouge water, where its preserves was not less abarming			-	- 1		Treate cheford and burded, he	dramer, d. c. May 11, 14, henrill at 11 -	ACCOMPANY A STANDARD TO A		
to the amendon, then destructive to the profiles;	Fages, Jeppes		30	30	230	I tubuctary Sales notices	Day and Back Tarren, Kingdom role H. D.	 Alternational and a second seco		 Control and the second s
mus we have moves beard of a more susprising						(Namona los Taurist and unate	Mr. Walander, Hall	ANTIVAS HEREARD		The second second second second
everype, these that of the accord and visbands dispersed						efections. This School is concentral	TATE halo of Linguage in the same of		2	A CONTRACTOR OF A CONTRACTOR O
in a boune, of which altmost everyroous lesars testi-	Back King-stepet	- 21	****	1 21	- 64	the the hometadomou of a docto	further been, of Lottrpech, as the enough of		<u>8</u>	
severy of the ravages of the anevertralable element.						indicated	tenere, province-mearinged, d. e.; May 17	A REAL PROPERTY AND A REAL PROPERTY OF	2	 Manual Annual Annua Annual Annual Annu
It appears probably, from an extenination of the	National Science, Grashy awa	294	119	aire 2		Volumentary Subscription and first	Man by st s, at the George Issa, Liverpool.		Telefort and the second and the seco	Statement of the second s
opertures which the fluid has made, and the direc-	Bulton street, Salperi	200	110	1002	600	S Committy resocreption, and Con-	We all a straining Liverpool.			CHINES COUNTS
tion in which the bricks, timber, do: have been						Excloses at C hurches.	WART Jreeph, Into of Ranhvary, in the coust	3		LISEN / Contract o
forced, that, attracted by the iron railing is front		858	1961	1939	62430		Oxford, (but now a principly in the King's h			A state of the sta
of the house, it entered the wall up one side the		~~~	-		1.0110	()	person), becave, d. e.; May 5, 13, Jane	3 SP 0007		
door, where it has shattered the bricks, icen to	Diameters.				_		FE, at Guildhall, London. Atla, Means. F	× \ / /	2	and the second se
piores the wood and brick-work between the duse-	Anascatura,	484	-				and Monday, Bellers.			A DESCRIPTION OF THE PARTY OF T
pillar and the arch-way of the door; lifted the	Lautarrenau Neuore, Marchall-st.		- 22.5	1.1		Coloutary Subscription.	WHARTON Robert and Henry Wharton, of I			Charlenger Discont Andrew Charlenger
boards on the sops; shatk the familight to pipper.	Carrasay, Mosley-steret	1000				EA(Do dotto	Crosby, in the crunty of Laaraster, jointer	and the second	2	A PARAMETER CONTRACTOR OF THE PARAMETER
burning part of the frame, and leaving a black and	CATBRES	194	101	365	104	Ditto ditto	boune-compositors, late constants in trade,	TANY MERITS		The second s
on the paint-work; thence it paterd up through the	SUNDAY SCHOOLS.	1000	241	1001	644.4		1, 19, June 9, at 11, at the George lass		ALT TO STATE	
sech door, splitting the bricks and the stone at the	* Establishment.			parts 1	4004		verpool, Att. Mr. Hudgens, Liverpool.			
buttom of the middle window, the glass of which	Collectate Clerch, Shade Hill,	-	-	40.03			"ILLMOTT Daniel, of Princens-sigure, Bo			
was shattered to pieces, and the whole frame dis-	St. Augh. Back King-doord	201	300	400)			hithe, in the ensaty of Sourcey, master-mar-			A STATE OF A
ballerd and forced into the hunt. Over the win-	St. Mary's, Back South Parade			100			merryhand, d. c.; May 5, at 10, May 15, a		CH. STOCK OF THE STOCK OF THE STOCK OF THE	CALCULATE COMPACTOR OF CALCULATION
dow it forced, is its way to the roof, a large hele.	52. Paul's, Green streat,	150	110	221			June 9, at 10, at timbihall, London.	The Areas of the blue; red, & black wedge	e are each measured from	
above which the suid appears exactly as if fames	Turnen-street	100		12211			Monars. Patternun and Polic, Old Broad-stra	the united as the common contra-		
had itsued from it. Its source appears next to kare	Jerney-stead,	314	-	100	- 1		DIVIDENDS.	the centre as the namman perter.	2	
been towards the chimacy: the case were shatteed	St. George's, St. George's	111	114	222			May EL. Marsair & Atkinson, London, murch	The blue wedges measured from the centre of	the arole more and	
to perces; the nidge stoors displaced; manay bricks	No. John's, St. John's street.	1141	100	941	- 1		22. J. Davies, Showabary, faz spinner,	for most the deal for the set of the		
and much consent form from the wall; and the lead	34. James's, St. James's street	100	100				25. W. Benky, Matchester, tanor.	THE AFER GRO MORPHANE IF AND STRUCTURES OF MIL	agable symplecal seases, ite	
in mony places forced up. It probably reached the	St. Michael's, Miller-treat.	204	314	22.1	61023		22. W. and A. Copp, Easter, danpers.	red wedges measured from the centre the	deaths from mounds & the	State of the state
rooms below by the chinacys. In the lower room	St. Peter's, Jackma's row.		1.00	100			22. W. Burnt, Exctor, despor and tailor		Automatic	10
the staces, plaster and paper are is several places	Algest Town		120	1201		6 Million	29. J. Dobell, Stapleburst, Keet, despar & to	black wedget measured from the centre b	he deaths from all other causes	V36 M30 S
broken, and the flaid, as if scatching its may out,	St. Clement's and St. Lake's,	~			- 1	Then to, perturps, the hargest behaved	22. J. Lor, London, watchouseman,	The Mark line agrees the red triangle in Me	This marks the boundary	and the states
scorched the gibling of the chimney glass, and	Beneret street	-	and a	الضما		in the Kaspeon. It call about	24. T. Cassidy, Liverpool, feather-movel	a tree statute into his ner oto vite o mongae provis		(1) 2-1 (2) (3) (3) (3)
peeled the top ornaments, but did not distark the	St. Stephen's, Bloom street		000	1000	******	<22,300, of which 2512 0 105	22. J. Williams, London, drapar,	of the deaths from all other causes during	the month.	
polished fire-iroon, just below. No squares of glass	Oldirid road	100	204	3.6		was contributed in small runn by	Th. M. B. Schlesinger London, indige-	In Rotakin all & down? all the black and	a compile with the set	and the second sec
were driven out in this room. In the room above,	Trinity, King's Head Yard.		200	2221		the Tenchers and Scholash.	chaol,	The manuer more, a April 1855, Die Mark Bree	C GOTTANDS INTER THE TERS	
in our corner, stood a bundle of rods, to which it	Hulme, Daks-staret	141	100	000			June 5. Ryder& Nasseyth, London, sugar-refi	in January & Pehruary 1856 the blue o	encides with the Mack	and the second se
made its way, pechaps from the chimney, between	All-Saints, Oxford road	12	101	201	-		BIBIOLCTIESS OF PARTNERSES.	72	44	VONVO
the lathing and the wall, as it forced off the planter,	Ardwick	100	122	001	34		John Harrison and Brothers, Manchrater, or	The entire areas may be compared by follow	wing the bare; the rea to the	
and sher a quantity of it against a chost of draw-			110	179	24		spinners Widow Welsh and Som, Manches	Hack lines melanna than		
on chosen or tacker for distant evidently with							common excitor Can Benedin	the second second second second		

Source:

Source:

All Broken of Locate States

A TA A CONTRACT CONTRACT PRODUCT TO A CONTRACT CONTRACT OF THE OWNER TO A CONTRACT CONTRACT OF THE OWNER TO A CONTRACT CONTRACT OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OF THE OWNER OF THE OWNER TO A CONTRACT OF THE OWNER OWNER OWNER OWNER TO A CONTRACT OF THE OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER OWNER OWNER TO A CONTRACT OWNER OWNER OWNER OWNER OWNER OWNER TO A CONTRACT OWNER TO A CONTRACT OWNER OW

Machine Learning



ARTIFICIAL INTELLIGENCE A technique which enables machines

to mimic human behaviour

MACHINE LEARNING

Subset of AI technique which use statistical methods to enable machines to improve with experience

DEEP LEARNING

Subset of ML which make the computation of multi-layer neural network feasible

Source : i Stack



Algorithms

- Algorithms- Black box
- Access by experts
- Demand expertise

Limit in participation





- Based on unstructured data
- Press releases: 2015- 2020 (October)

Analysis

Αθήνα, 27 Ιουνίου 2015

ΔΕΛΤΙΟ ΤΥΠΟΥ

Συνελήφθησαν πέντε (5) άτομα για διακίνηση ναρκωτικών

Κατασχέθηκαν ποσότητες κατεργασμένης και ακατέργαστης κάνναβης, κοκαΐνης, ηρωίνης και ναρκωτικών δισκίων

Συνελήφθησαν, στις 23-6-2015 στον Άγιο Δημήτριο και στην Κορινθία, από την Υποδιεύθυνση Δίωξης Ναρκωτικών της Διεύθυνσης Ασφάλειας Αττικής, τέσσερις (4) ημεδαποί ηλικίας 43,48,51 και 58 ετών και μία 53χρονη αλλοδαπή υπήκοος Αλβανίας, για διακίνηση ναρκωτικών ουσιών σε περιοχές της Αττικής και της περιφέρειας.

Σε βάρος των συλληφθέντων σχηματίσθηκε ποινική δικογραφία, για τις κατά περίπτωση παραβάσεις των νόμων περί ναρκωτικών, περί όπλων και περί αλλοδαπών.

Analysis

Date	Date_2 🔺 Day	Day_2	Team	Native	N_Arrested	Foreign	F_Arrested	Nationality	Nationality_2	Age	Min	Max	Sex	Location	1
12/9/2015	12/9/2015 Σάββατο	6	0			1		Σερβία	1	34			1	Διεθνής Αερολιμένα	1
25/9/2013	25/9/2013 Τετάρτη	3	0	()			Ελληνική	0	19			1		
25/9/2013	25/9/2013 Τετάρτη	3	0	()			Ελληνική	0	20			1		
25/9/2013	25/9/2013 Τετάρτη	3	0	()			Ελληνική	0	24			1		
25/9/2013	25/9/2013 Τετάρτη	3	0	()			Ελληνική	0	30			1		
25/9/2013	25/9/2013 Κυριακή	7	0	()			Ελληνική	0	26			1		
25/9/2013	25/9/2013 Κυριακή	7	0	()			Ελληνική	0	36			1		
25/9/2013	25/9/2013 Τετάρτη	3	0			1		Αλγερία	1	32			1		
22/2/2015	22/2/2015 Παρασκευή-Σάββατο	5	2			1	. 4	Ι Γεωργία	1	0			2		
22/2/2015	22/2/2015 Παρασκευή-Σάββατο	5	0			1		Βουλγαρία	1	0			2		
30/9/2015	30/9/2015 Παρασκευή-Σάββατο	5	0			1		Βουλγαρία	1	0			2		
30-31/1/2015	30/1/2015 Σάββατο	6	0			1		Αφγανιστάν	1	34			1	Πλατεία Βικτωρίας	1
30-31/1/2015	30/1/2015 Σάββατο	6	1			1		Νέα Γουινέα	1	0			1	Περιοχή Κυψέλης	1
30-31/1/2015	30/1/2015 Σάββατο	6	1			1		Σιέρα Λεόνε	1	0			1	Περιοχή Κυψέλης	1
3/10/2015	3/10/2015 Σάββατο-Κυριακή	6	2			1	. 2	2 Γεωργία	1	0			2	Νοσοκομείο	
5/2/2015	5/2/2015 Πέμπτη	4	0			1		Βουλγαρία	1	0			2	Νοσοκομείο	
5/2/2015	5/2/2015 Πέμπτη	4	0			1		Ρωσία	1	0			2	Νοσοκομείο	
7-8/2/2015	7/2/2015 Σάββατο-Κυριακή	6	0			1		Αλβανία	1	0			2	Νοσοκομείο	



Percentage of arrests by sex

Men 📕 Women 📕 Not defined



Mean age of arrests by sex

Men 📕 Women 📕 Not defined



Data exploration

Native Foreigner

Frequency of arrests by Nationality



Data exploration

Total arrests per day of week







3 Classes model I:

Algorithm: Decision tree

Metrics:

- F1_score \rightarrow 0.69
- Accuracy \rightarrow 0.695

<u>3 Classes model II:</u>

Algorithm: Random Forest

Metrics:

• F1_score \rightarrow 0.69

• Accuracy \rightarrow 0.695



Models

2 Classes model I:

Algorithm: Decision tree

Metrics:

- F1_score \rightarrow 0.74
- Accuracy \rightarrow 0.74

2 Classes model II:

Algorithm: Random forest

Metrics:

- F1_score \rightarrow 0.75
- Accuracy \rightarrow 0.752



Variables importance

3 Classes model I:

- Age
- Day of week
- Total Retirees
- Postal code
- Population

3 Classes model II:

- Age
- Day of week
- Postal code
- Total employees
- Nationality



Variables importance

2 Classes model I:

- Age
- Total employees
- Day of week
- Postal code
- Nationality

2 Classes model II:

- Age
- Day of week
- Postal code
- Total employees
- Nationality

Minor Importance Variables

DECISION TREE (3 CLASSES)

Native/Foreigner

Sex

RANDOM FOREST (3 CLASSES)

Sex

Native/Foreigner

DECISION TREE (2 CLASSES)

Sex

Team **or** not

RANDOM FOREST (2 CLASSES)

Native/Foreigner

Sex





- Accuracy of prediction: 0.752
- F1→ 0.75
- Variable importance:
- 1. Age
- 2. Day of the week
- 3. Postal code
- 4. Total employees
- 5. Nationality



Source : <u>Frontiersin</u>





Minor importance variables:

1. Native / Foreigners

2. Sex

Conclusion

- Our purpose fulfilled
- Good accuracy of model
- Further investigation required



- Advantages of unstructured data:
- 1. Better policing
- 2. More efficient Human Resource Management
- From experiment to action



Source: Unsplash



Thank you!



marilouvardaki@yahoo.gr https://www.linkedin.com/in/maria-eleni-vardaki