

Χρήση Machine Learning για βελτιστοποίηση Software Project Schedule Management & Project Cost Estimation: Case Study

Διπλωματική Εργασία του Κοτζαμπάση Στυλιανού
Επιβλέπων Καθηγητής: Μαντάς Μιχαήλ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΤΜΗΜΑΤΟΣ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ



Περιεχόμενα

1. Εισαγωγή
2. Βιβλιογραφική Ανασκόπηση
3. Μεθοδολογία
4. Αποτελέσματα
5. Συμπεράσματα



Εισαγωγή



Πρόβλημα



Για αποτελεσματικό Project Management, χρειάζεται **γνώση των πραγματικών δεδομένων** που σχετίζονται με τις **εργασίες**, τους **πόρους** και γενικότερα το σύνολο των **Projects** (Kanakaris et al., 2020).

Η διαχείριση του **χρονοπρογραμματισμού** και του **κόστους των έργων λογισμικού**, αποτελούν δύο από τις κορυφαίες προτεραιότητες στη βιομηχανία παραγωγής προϊόντων λογισμικού (Rana and Wei, 2020).

Σύμφωνα με τους αναλυτές της Gartner, **το 80% των εργασιών** (tasks) του Project Management του σήμερα, θα πραγματοποιείται από συστήματα τεχνητής νοημοσύνης (AI) που αξιοποιούν την μηχανική μάθηση (machine learning) **μέχρι το 2030** (Chennouk et al., 2020).

Οι **μικρές** και **μικρομεσαίες επιχειρήσεις** συνεχίζουν να χρησιμοποιούν **απαρχαιωμένες μεθόδους** που οδηγούν σε συχνές καθυστερήσεις και αύξηση του κόστους του έργου (Wei and Rana, 2019).



Σκοπός



Διερεύνηση χρήσης Machine Learning στις περιοχές γνώσης του Software Project Schedule Management & του Project Cost estimation.

- Δημιουργία 2 λύσεων

Η 1η: Προβλέπει το χρόνο υλοποίησης μιας ενέργειας, βάσει των παρελθοντικών δεδομένων και πόρου που την αναλαμβάνει, με χρήση regression models.

Η 2η: Κατανέμει τις εργασίες στους πόρους, αξιοποιώντας την παραπάνω γνώση.

Σύγκριση προσδοκόμενων αποτελεσμάτων με πραγματικά, για την εξαγωγή συμπερασμάτων σχετικά με την ακρίβεια των λύσεων και την αξιοποίηση τους σε πραγματικές συνθήκες.



Συνεισφορά



Έρευνες προσεγγίζουν την χρήση του Machine Learning και συγκεκριμένα των Regression Models στην πρόβλεψη του χρόνου υλοποίησης των Software Projects (Finnie et al., 1997; de Barcelos Tronto et al., 2007; Nassif et al., 2013).

Σε ερευνητικό επίπεδο συνεισφέρει :

- Στον έλεγχο, την αξιολόγηση και την σύγκριση των αποδοτικότερων μοντέλων παλινδρόμησης για την πρόβλεψη του χρόνου υλοποίησης ενός software project.
- Στην ανάδειξη των εμποδίων που προκύπτουν με την χρήση της μεθοδολογίας waterfall στο software project management.

Σε επαγγελματικό επίπεδο προσφέρει:

- Στις μικρές και τις μικρομεσαίες επιχειρήσεις μία βάση για να δομήσουν λύσεις που αξιοποιούν το machine learning για την πρόβλεψη του μελλοντικού χρόνου υλοποίησης σε software projects, καθώς και για την κατανομή των ενεργειών στους πόρους.



Βιβλιογραφική Ανασκόπηση





Διαδικασία Αναζήτησης Άρθρων

1η Αναζήτηση:

Keywords: "Project Management" & "AI"

Έτη: 2017-2020

Fields: Engineering, Computer Science, Decision Sciences, Business-Management & Accounting

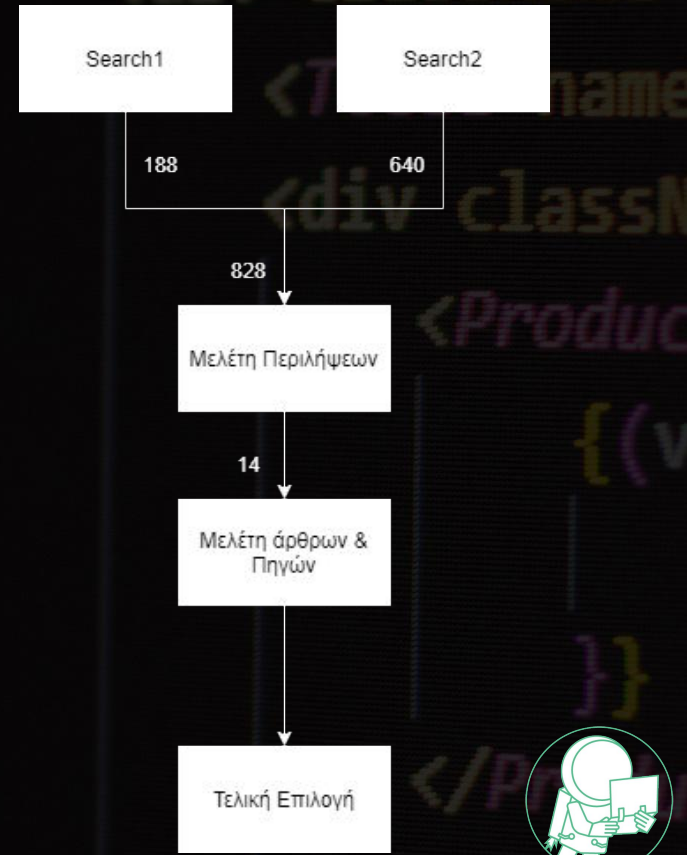
2η Αναζήτηση:

Keywords "Project Management" & "Machine Learning"

Έτη: 2016-2020

Fields: Engineering, Computer Science, Decision Sciences, Business-Management & Accounting

14 → 10 → 31





Project Time Management & Project Cost Management

Οι 47 διαδικασίες της διαχείρισης έργου κατηγοριοποιούνται επιπλέον σε 10 περιοχές γνώσης. 2 από τις σημαντικότερες είναι (PMI, 2013):

- 1) **Η διαχείριση χρόνου έργου** περιλαμβάνει όλες τις διαδικασίες που απαιτούνται για την έγκαιρη ολοκλήρωση του.
- 2) **Η διαχείριση κόστους έργου** περιλαμβάνει τις διαδικασίες που εμπλέκονται στο σχεδιασμό, την εκτίμηση, τον προϋπολογισμό, τη χρηματοδότηση, τη διαχείριση και τον έλεγχο του κόστους.



Μεθοδολογία Waterfall

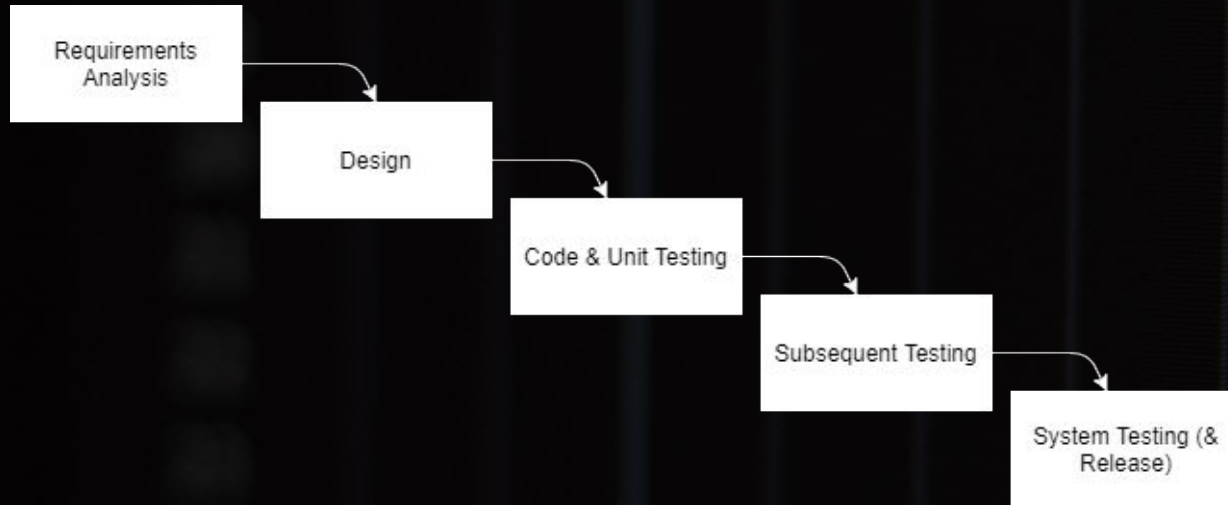
Δημιουργία: Dr. Winston Royce to 1970

Σκοπός: Συμβολή στην ανάπτυξη λογισμικού (McCormick, 2012).

Βελτίωση: Dr. Barry Boehm 1974 -1976 (Charvat, 2003).

Σήμερα: Πιο διαδεδομένη μέθοδος στο Software Development (McCormick, 2012).

Φάσεις που προτείνονται (McCormick, 2012):



Δυσκολίες στο Project Schedule Management



1. Ο χρόνος ολοκλήρωσης του Project μπορεί να υπερβαίνει από 25% έως και 100% την αρχική πρόβλεψη (Paratheocharous et al., 2017).
2. Έλλειψη σωστού χρονοπρογραμματισμό του έργου από τους PMs, ώστε να επιτευχθούν έγκαιρα οι στόχοι της ομάδας (Englund and Graham, 2019).

Έτσι, τόσο η προβλεπόμενη ημερομηνία παράδοσης, όσο και η πρόβλεψη του κόστους του έργου, **τείνουν να είναι μη ρεαλιστικές.**



Μέθοδοι Υπολογισμού Χρόνου Υλοποίησης



Οι μέθοδοι υπολογισμού του χρόνου υλοποίησης ενός Project μπορούν να καταταχθούν σε 3 βασικές κατηγορίες (Trendowicz και Jeffery, 2014):

- 1) τις data-driven,
- 2) τις expert-based και
- 3) τις υβριδικές.

Οι expert-based μέθοδοι είναι ευρέως διαδεδομένοι και λόγω έλλειψης αποδείξεων για την δυνατότητα εκτίμησης των data-driven μεθόδων (Molokken & Jorgensen, 2003).

Μελέτη **304 δημοσιεύσεων** σε **76 επιστημονικά περιοδικά** (Jorgensen και Shepperd, 2006).



Ανάγκη συλλογής αποδείξεων για την αποτελεσματικότητα των data-driven μεθόδων, με στόχο την καλύτερη εκτίμηση του χρόνου υλοποίησης.





Data Driven Μέθοδοι

Memory-based μέθοδοι

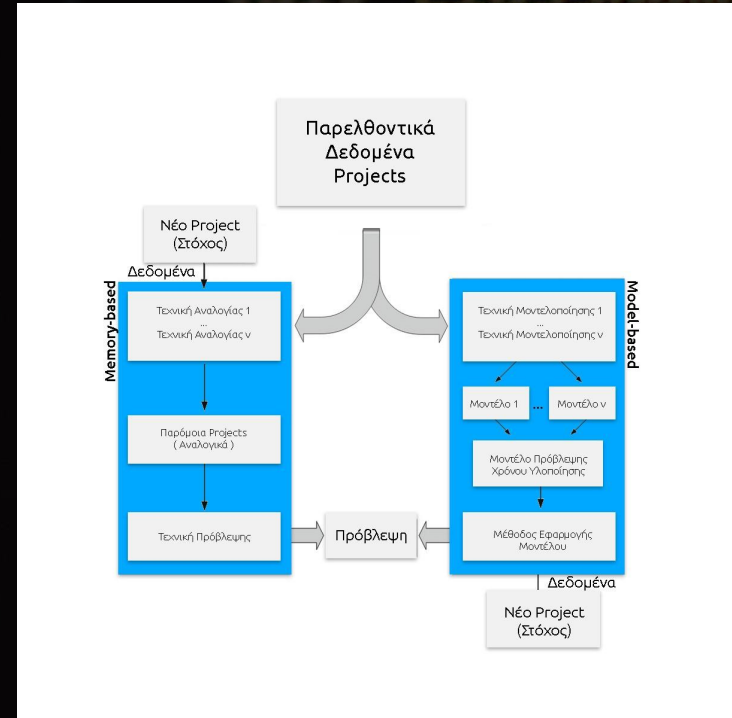
Αρκετές έρευνες, από το 1993 έως και το 1997, με αισιόδοξα αποτελέσματα (MacDonell και Shepperd, 2003).

Η αισιοδοξία ήταν αβάσιμη (Jorgensen και Shepperd, 2006).

Model-based μέθοδοι

Έρευνες για την χρήση των neural networks & των regression models (Finnie et al., 1997; de Barcelos Tronto et al., 2007; Attarzadeh et al., 2012; Nassif et al., 2013;), με **αισιόδοξα αποτελέσματα**.

Από τις 304 δημοσιεύσεις - το **49%** έχει αξιοποιήσει τα **regression models** - το **3% Case Studies** (Jorgensen και Shepperd, 2006).



Χρήση Machine Learning για Βελτίωση Software Project Schedule & Cost Management



Η χρήση της τεχνολογίας του Machine Learning μπορεί να συμβάλει:

- 1) Στο **ρεαλιστικό προϋπολογισμό του χρόνου** που απαιτείται για την υλοποίηση μιας ενέργειας (task) (Hoda and Murugesan, 2016).
- 2) Κατ' επέκταση, στον **ορθότερο καταμερισμό των ενεργειών** (tasks) στους πόρους της ομάδας, βάσει των χαρακτηριστικών τους.



Regression Models



Η παλινδρόμηση είναι μια τεχνική για την μοντελοποίηση και την ανάλυση των συσχετίσεων που υπάρχουν μεταξύ των μεταβλητών, καθώς και πως αυτές επηρεάζουν το τελικό αποτέλεσμα (Rana & Wei, 2020).

Υπάρχουν αρκετά μοντέλα παλινδρόμησης:

- 1) Linear Regression
- 2) Multiple Linear Regression
- 3) Decision Tree
- 4) Random Forest Regression
- 5) Polynomial Regression

Εφαρμογή στο Excel



Χρήση Regression Models για πρόβλεψη Χρόνου Υλοποίησης



Έρευνες με χρήση παλινδρόμησης για την πρόβλεψη της προσπάθειας που θα χρειαστεί για την υλοποίηση ενός Project (Finnie et al., 1997; de Barcelos Tronto et al., 2007; Nassif et al., 2013).

Η προσπάθεια = εργατοώρες ή κόστος

Μελέτη Nassif et al. (2013), φαίνεται πως τα μοντέλα παλινδρόμησης που ερευνήθηκαν είχαν αρκετά μεγάλη ακρίβεια, **με το μοντέλο random forest regression** να ξεχωρίζει σε σύγκριση με τα μοντέλα decision tree και multiple linear regression.



Μεθοδολογία



Σχεδιασμός της έρευνας



Πραγματοποιήθηκε αξιοποιώντας τη μεθοδολογία των Rana & Wei (2020) και τις αρχές των Runeson & Höst (2009) για την δημιουργία έρευνας τύπου case study για software engineering. Χωρίζεται σε 7 στάδια:

1. Συλλογή & Επεξεργασία Δεδομένων,
2. Ανάλυση Δεδομένων,
3. Δημιουργία Μοντέλου Πρόβλεψης,
4. Παραγωγή Δεδομένων Πρόβλεψης,
5. Δημιουργία Προγράμματος Κατανομής Εργασιών,
6. Δημιουργία Δεδομένων Πρόβλεψης & Αξιολόγηση και
7. Σύγκριση Δεδομένων με Πραγματικά Αποτελέσματα.

Φάσεις	Στάδια
Προετοιμασία Δεδομένων	1
Εκτέλεση Ενεργειών	2,3,4,5
Ανάλυση	6
Εξαγωγή Συμπερασμάτων	7

Runeson & Höst (2009)





Συλλογή & Επεξεργασία Δεδομένων

Συλλέχθηκαν ποιοτικά και ποσοτικά δεδομένα για 8 πόρους, 110 Projects και 3830 Tasks.
60 Projects Ecommerce & 50 Projects Brand Websites.

Χαρακτηριστικά Πόρων

ResourceID	ResourceCost (Annual)	ResourceCost(Hourly)	TAG	SkillLevel	YOE
TDS	29049	15,71915584	WebDeveloper	80	6
FLD	24392	13,1991342	WebDeveloper	50	3
XDI	20038	10,84307359	WebDeveloper	30	1
GEO	20038	10,84307359	WebDesigner	60	3

Λίστα Tasks

Phase	TaskName	TaskID	Res_Tag	Project_Type
Requirements Phase	Συμπλήρωση ερωτηματολογίου από πελάτη	REQ1	ProjectManager	Ecommerce, BrandWebsite
Requirements Phase	Έλεγχος από ομάδα ανάπτυξης & βελτιστοποιήσεις	REQ2	ProjectManager	Ecommerce, BrandWebsite
Requirements Phase	Αποστολή τελικού ερωτηματολογίου στον πελάτη και τελική έγκριση	REQ3	ProjectManager	Ecommerce, BrandWebsite



Συλλογή & Επεξεργασία Δεδομένων



Παρελθοντικά Δεδομένα

ProjectID	Project_Type	Phase	TaskName	TaskID	Res_Tag	ResourceID	Time_Spent
ES0001	Ecommerce	Requirements Phase	Συμπλήρωση ερωτηματολογίου από πελάτη	REQ1	ProjectManager	SGK	30
ES0001	Ecommerce	Requirements Phase	Έλεγχος από ομάδα ανάπτυξης & βελτιστοποιήσεις	REQ2	ProjectManager	SGK	50
ES0001	Ecommerce	Requirements Phase	Αποστολή τελικού ερωτηματολογίου στον πελάτη και τελική έγκριση	REQ3	ProjectManager	SGK	15



Ανάλυση Δεδομένων



Ανάλυση Δεδομένων Πόρων

TaskID	AvailableResources	AverageTotal	MaxTotal	MinTotal
REQ1	SGK	43	65	25
DES3	GEO	421	475	335
DEV5	TDS, FLD, XDI	101	155	50
DEV10	SGI, JMK, JIG	328	450	225

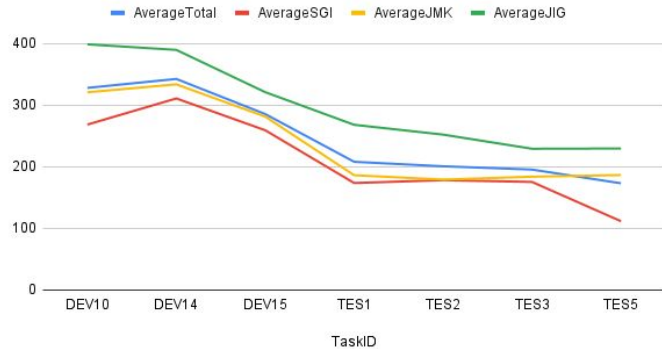
Στα tasks που εμπλέκονται περισσότεροι πόροι η διαφορά min & max μπορεί να είναι **>400%**.



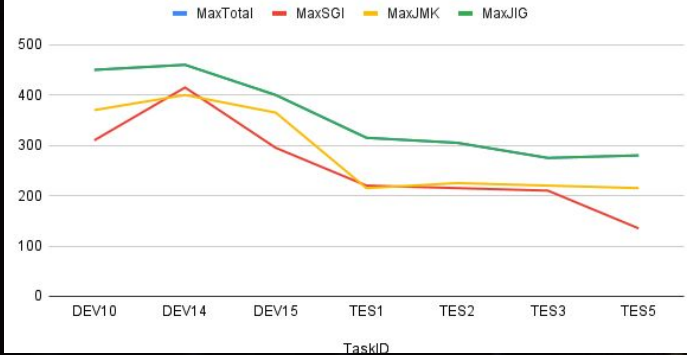
Ανάλυση Δεδομένων

Για τους WebConsultants:

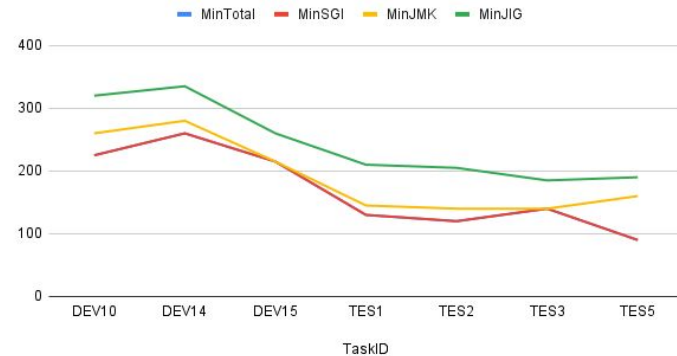
Average Time per Task



Max Time Per Task



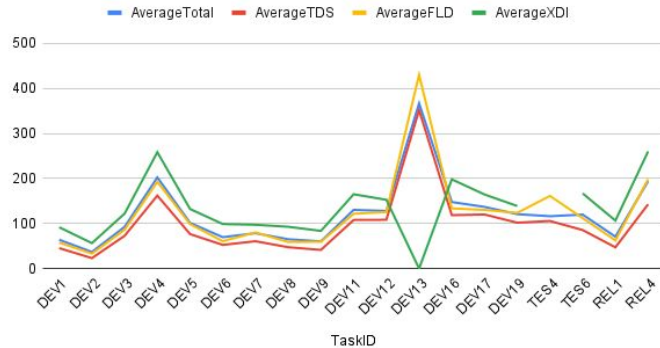
Min Time Per Task



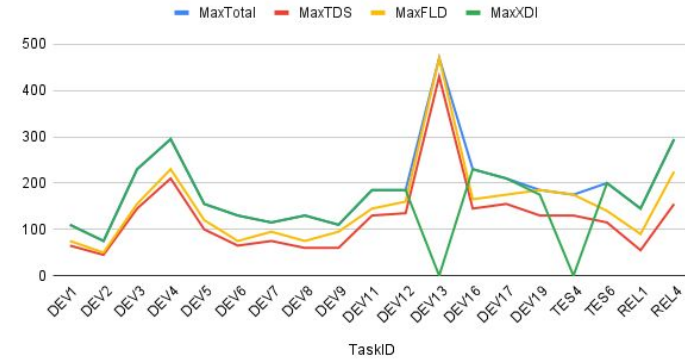
Ανάλυση Δεδομένων

Για τους WebDevelopers:

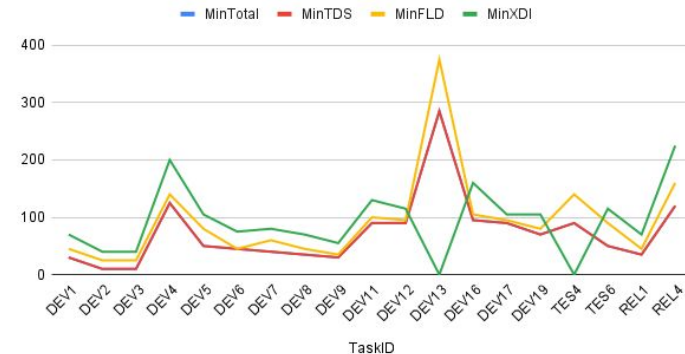
Average Time Per Task



Max Time Per Task



Min Time per Task





Δημιουργία Μοντέλου Πρόβλεψης

Βάσει των ερευνών των Nassif et al. (2013), των Wauters και Vanhoucke (2014) και των Maulud και Abdulazeez (2020), διερευνώνται και αξιολογούνται 5 μοντέλα παλινδρόμησης, τα οποία φαίνεται να είχαν αρκετά μεγάλη ακρίβεια (>80%) στις προηγούμενες έρευνες:

1. Decision Tree Regression
2. Random Forest Regression
3. Multiple Linear Regression
4. Polynomial Regression
5. Support-Vector Machine Regression

Για την αξιολόγηση τους, χρησιμοποιείται η πλατφόρμα του Google Collab και υλοποιείται η παρακάτω διαδικασία για το καθένα από αυτά.





Δημιουργία Μοντέλου Πρόβλεψης

Εισαγωγή βιβλιοθηκών

```
[ ] import numpy as np
import pandas as pd
```

Εισαγωγή δεδομένων με αριθμητική σειρά

```
▶ dataset = pd.read_csv('DataSK.csv')
X = dataset.iloc[:, :-1]
y = dataset.iloc[:, -1]
```

Μετατροπή δεδομένων σε διανύσματα

```
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
ct = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), [0,1])], remainder='passthrough')
X = ct.fit_transform(X)
```

Διαχωρισμός δεδομένων

```
▶ from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
```

Δεδομένα - Input

TaskID	TAG	SkillLevel	YOE	ResourceCost(Annual)	Time_Spent
REQ1	ProjectManager	70	4	24392	30
REQ2	ProjectManager	70	4	24392	50





Δημιουργία Μοντέλου Πρόβλεψης

Εκπαίδευση Μοντέλου

```
▶ from sklearn.tree import DecisionTreeRegressor  
regressor = DecisionTreeRegressor(random_state = 0)  
regressor.fit(X_train, y_train)
```

Παραγωγή μελλοντικών τιμών

```
▶ y_pred = regressor.predict(X_test)  
np.set_printoptions(precision=2)  
s2 = y_test.to_numpy()  
print(np.concatenate((y_pred.reshape(len(y_pred),1), s2.reshape(len(y_test),1)),1))
```

```
👤 [[ 56.9   55.   ]  
    [ 72.1   30.   ]  
    [ 60.67  65.   ]  
    ...  
    [109.55  90.   ]  
    [ 46.43  55.   ]  
    [ 41.86  40.   ]]
```



Δημιουργία Μοντέλου Πρόβλεψης



Σύγκριση & αξιολόγηση μοντέλου

```
[ ] from sklearn.metrics import r2_score  
    r2_score(y_test, y_pred)
```

```
0.9554455129874584
```

Αξιολόγηση Regression Models

Regression Model	Accuracy Level
Random Forest	95,55%
Decision Tree	95,54%
Polynomial Regression	95,54%
Multiple Linear Regression	94,94%
Support-Vector Machine Regression	89,31%





Παραγωγή Δεδομένων Πρόβλεψης

Πιθανοί Συνδυασμοί Πόρων-Tasks (94 Συνδυασμοί)

ResourceID	TaskID	Res_Tag	SkillLevel	YOE	ResourceCost
SGK	REQ1	ProjectManager	70	4	24392
SGK	REQ2	ProjectManager	70	4	24392
SGK	REQ3	ProjectManager	70	4	24392
SGK	REQ4	ProjectManager	70	4	24392
SGK	DES1	ProjectManager	70	4	24392
SGK	DES2	ProjectManager	70	4	24392

Προβλέψεις χρόνου υλοποίησης ανά συνδυασμό

TaskID	ResourceID	Time_Spend_Prediction
REQ1	SGK	43
REQ2	SGK	47
REQ3	SGK	27
REQ4	SGK	56
DES1	SGK	43
DES2	SGK	



Δημιουργία Προγράμματος Κατανομής Εργασιών



Ποροκεντρική προσέγγιση γιατί:

1. Οι πόροι έχουν μεγαλύτερη διάρκεια και σταθερότητα από τα Projects.
2. Μεγάλη Επαναληψιμότητα των ενεργειών μεταξύ των Projects.
3. Βασικός στόχος είναι η μέγιστη αξιοποίηση των πόρων και κατ' επέκταση η υλοποίηση όσο των δυνατών περισσότερων projects παράλληλα.

Άρα, το πρόγραμμα θα κατανέμει Tasks στους πόρους και όχι το αντίστροφο.

Βάσει των παραπάνω και της έρευνας του Light (2009), επιλέχθηκε η αξιοποίηση του waterfall μοντέλου.



Δημιουργία Προγράμματος Κατανομής Εργασιών



Βάση προδιαγραφών σχεδιάστηκε λύση σε python, η οποία έχοντας ως default parameters:

- 1) Τα tasks και τους χρόνους ανά πόρο,
- 2) τα default phases που αποτελούν μια λίστα των tasks και
- 3) τους πόρους.

Δέχεται ως input Projects που αποτελούνται από default phases & κατανέμει τα tasks στους αντίστοιχους πόρους για κάθε ημέρα.

Η κατανομή των tasks στους πόρους **γίνεται με βάση το tag** και ως διαθέσιμος χρόνος για κάθε πόρο **θεωρείται το 8ωρο/ημέρα**.



Δημιουργία Προγράμματος Κατανομής Εργασιών



Output Προγράμματος Κατανομής Εργασιών

day	assignment_id	resource_id	task_id	project_id	estimated_time
1	0	SGK	REQ1	1	0:43:00
1	1	SGK	REQ2	1	0:47:00
1	2	SGK	REQ3	1	0:27:00



Δημιουργία Δεδομένων Πρόβλεψης



Phase	Tasks
1	REQ1,REQ2,REQ3,REQ4
2	DES1, DES2, DES3, DES4, DES5, DES6
3	DEV1, DEV2, DEV3, DEV4, DEV6, DEV10, DEV11, DEV18, DEV19, DEV20
4	DEV1, DEV2, DEV3, DEV4, DEV6, DEV7, DEV8, DEV9, DEV10, DEV11, DEV18, DEV19, DEV20
5	TES1, TES3, TES6, TES7
6	TES1, TES2, TES3, TES6, TES7
7	REL1, REL5

Συνδυασμός Phases 1,2,3,5,7 => Brand Website

Συνδυασμός Phases 1,2,4,6,7 => Ecommerce



Δημιουργία Δεδομένων Πρόβλεψης



Προσθέτουμε ως Input 3 Brand Websites & 3 Ecommerce Projects



day	assignment_id	resource_id	task_id	project_id	estimated_time
1	0	SGK	REQ1	1	0:43:00
1	1	SGK	REQ2	1	0:47:00
1	2	SGK	REQ3	1	0:27:00
1	3	SGK	REQ4	1	0:56:00
1	4	SGK	DES1	1	0:43:00



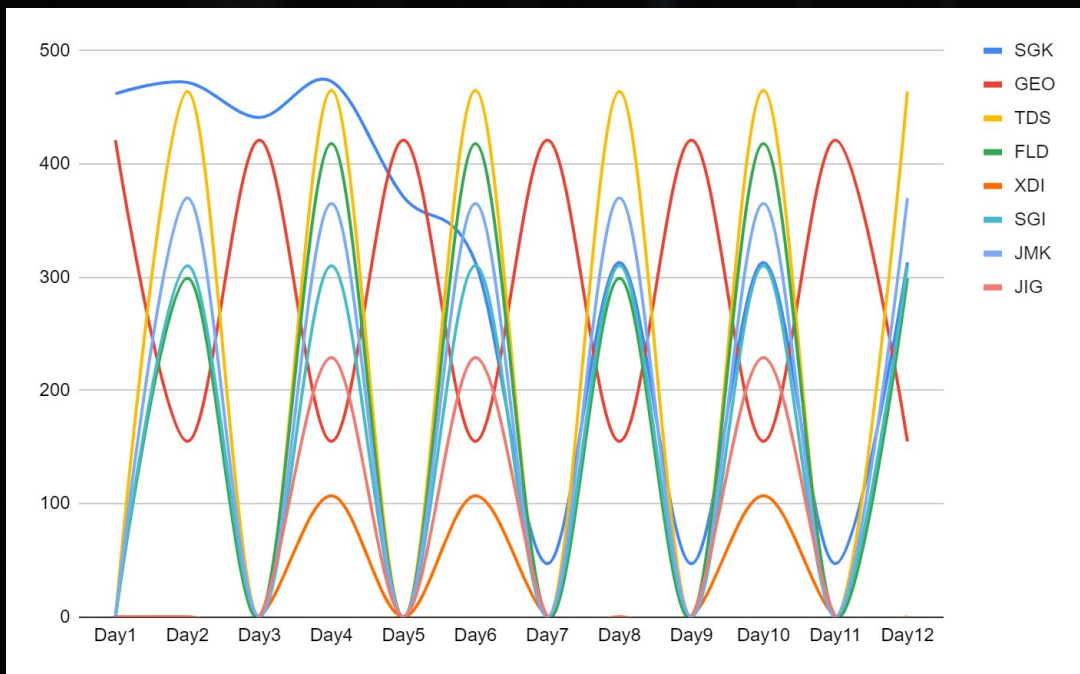
Αποτελέσματα



Αξιολόγηση προγράμματος κατανομής εργασιών



Χρόνος υλοποίησης ανά πόρο ανά ημέρα



Αξιολόγηση προγράμματος κατανομής εργασιών



Χρόνος υλοποίησης ανά τύπο πόρου

ResTag	AvailableResources	Estimated_Time	Est_Time_Per_Resource
PROJECT_MANAGER	1	3612	3612
WEB_DESIGNER	1	3456	3456
WEB_CONSULTANT	3	4572	1524
WEB_DEVELOPER	3	5259	1753



Σύγκριση δεδομένων πρόβλεψης με πραγματικά αποτελέσματα



Project	TotalTime_Est	TotalTime_Act	Per.Change
1	2621	2735	4,35%
2	3072	3090	0,59%
3	3072	3380	10,03%
4	2621	2720	3,78%
5	3072	3170	3,19%
6	2621	2805	7,02%

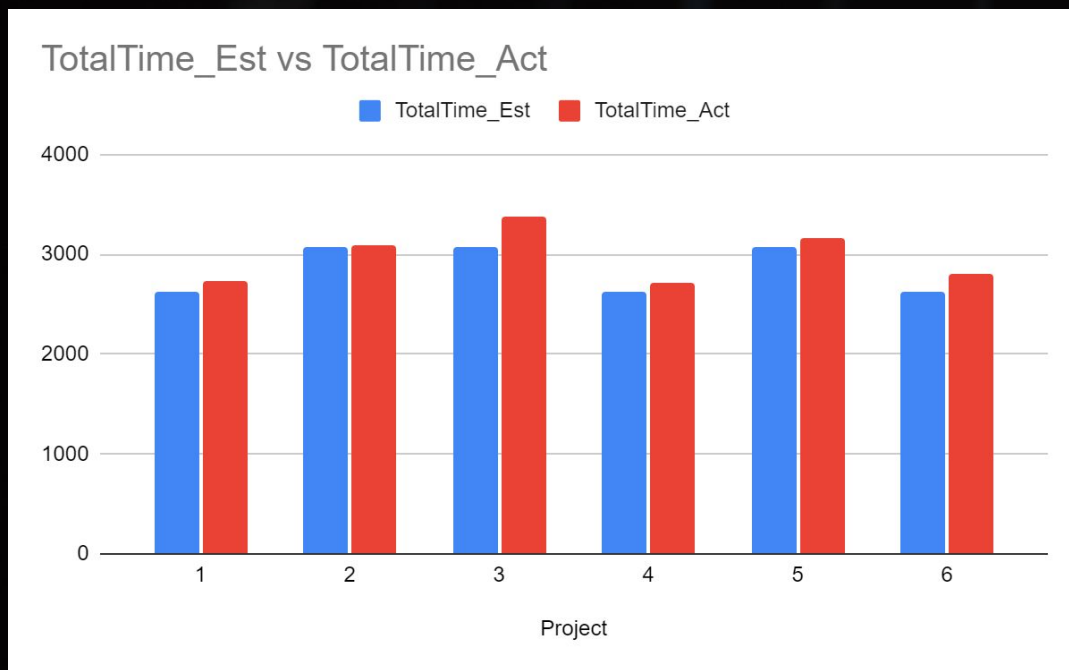
Σύγκριση Δεδομένων Χρόνου/Project



Σύγκριση δεδομένων πρόβλεψης με πραγματικά αποτελέσματα



Σύγκριση Δεδομένων Χρόνου/Project



Μ.Ο αποκλίσεων Projects → **4,82%**



Σύγκριση δεδομένων πρόβλεψης με πραγματικά αποτελέσματα



Σύγκριση Δεδομένων Χρόνου ανά πόρο

Resource	TotalTime_Est	TotalTime_Act	Per.Change
SGK	3612	3990	10,47%
GEO	3456	3510	1,56%
TDS	2787	2925	4,95%
FLD	2151	2295	6,69%
XDI	321	350	9,03%
SGI	1860	1935	4,03%
JMK	2205	2215	0,45%
JIG	687	680	-1,02%

Μ.Ο αποκλίσεων Πόρων→ **4,52%**



Συμπεράσματα



Συμπεράσματα



1η λύση: Συγκρίνοντας των μ.ο. των αποκλίσεων των Projects (4,82%) με το γεγονός πως ο χρόνος ολοκλήρωσης ενός Project μπορεί να υπερβαίνει από 25% έως και 100% την αρχική πρόβλεψη (Rana and Wei, 2020), καταλήγουμε στο συμπέρασμα πως το machine learning και συγκεκριμένα τα regression models, αποτελούν ένα χρήσιμο εργαλείο στην φαρέτρα του σύγχρονου Project Manager.

2η λύση: Φαίνεται πως είναι ιδιαίτερα σημαντικό να επιλεχθεί η κατάλληλη μεθοδολογία βάσει της ομάδας, των αναγκών και της φύσης των Projects που υλοποιούνται. Δεν υπάρχει καθολική λύση και δεν πρέπει οι προτεινόμενες μεθοδολογίες να εφαρμόζονται «τυφλά», καθώς κάθε οργανισμός είναι ιδιαίτερος.



Περιορισμοί της έρευνας

Περιορισμένος όγκος δεδομένων:

- Tasks
- Projects
- Πόρων



Μελλοντικές Επεκτάσεις



Συγκέντρωση μεγαλύτερου όγκου δεδομένων, επανέλεγχος των regression models και εκπαίδευση της 1ης λύσης βάσει αυτών.

Αξιοποίηση λύσεων για διαφορετικού τύπου Projects στο Software Development, π.χ. mobile applications.

Βελτίωση 2ης λύσης βασιζόμενοι σε waterfall μεθοδολογία με overlap ή σε agile μεθοδολογία, αποφεύγοντας την μεγάλη διαθεσιμότητα πόρων που δημιουργήθηκε.

Η επέκταση του υπάρχοντος data-driven μοντέλου εκτίμησης, προσθέτοντας χαρακτηριστικά των expert-based models, με σκοπό την δημιουργία ενός υβριδικού αποδοτικότερου μοντέλου.

Αξιοποίηση των λύσεων συνδυαστικά με άλλους παράγοντες, για αξιολόγηση & βελτιστοποίηση ομάδας, βάσει των Projects και της τάσης που δημιουργείται.



Ερωτήσεις;



Βιβλιογραφία



1. Attarzadeh, I. et al. (2012) "Proposing an enhanced artificial neural network prediction model to improve the accuracy in software effort estimation," in 2012 Fourth International Conference on Computational Intelligence, Communication Systems and Networks, pp. 167–172.
2. Charvat, J. (2003). Project management methodologies: Selecting, implementing, and supporting methodologies and processes for projects. John Wiley & Sons.
3. Chennouk, H et al. (2020) "The Machine Learning in Project Management: A Systematic Mapping Study", Journal of Theoretical and Applied Information Technology. 98(10), pp. 1550-1563
4. de Barcelos Tronto, I. F et al. (2007) "Comparison of artificial neural network and regression models in software effort estimation," in 2007 International Joint Conference on Neural Networks, pp. 771–776.
5. Englund, R. and Graham, R. J. (2019) Creating an environment for successful projects. Berrett-Koehler Publishers.
6. Finnie, G. R. At al. (1997) "A comparison of software effort estimation techniques: Using function points with neural networks, case-based reasoning and regression models," Journal of systems and software. Elsevier, 39(3), pp. 281–289.
7. Hoda, R. and Murugesan, L. K. (2016) "Multi-level agile project management challenges: A self-organizing team perspective," Journal of Systems and Software, 117, pp. 245–257.
8. Jorgensen, M. and Shepperd, M. (2006) "A systematic review of software development cost estimation studies," IEEE Transactions on software engineering. IEEE, 33(1), pp. 33–53.
9. Kanakaris, N. et al. (2020) "Combining Machine Learning and Operations Research Methods to Advance the Project Management Practice" in Parlier, G. H., Liberatore, F., and Demange, M. (eds.) Operations Research and Enterprise Systems. Cham: Springer International Publishing, pp. 135–155.
10. Light, M. (2009) "How the Waterfall Methodology Adapted and Whistled Past the Graveyard," Gartner Research.
11. M Wei, W. and Rana, M.E., (2019) "Software Project Schedule Management Using Machine Learning & Data Mining". International Journal of Scientific & Technology Research, 8(9), pp.1385-1389.
12. MacDonell, S. G. and Shepperd, M. J. (2003) "Combining techniques to optimize effort predictions in software project management," Journal of Systems and Software. Elsevier, 66(2), pp. 91–98.
13. Maulud, D. and Abdulazeez, A. M. (2020) "A Review on Linear Regression Comprehensive in Machine Learning," Journal of Applied Science and Technology Trends, 1(4), pp. 140–147.
14. McCormick, M. (2012) "Waterfall vs. Agile methodology," MPCs.
15. Molokken, K. and Jorgensen, M. (2003) "A review of software surveys on software effort estimation," in 2003 International Symposium on Empirical Software Engineering, 2003. ISESE 2003. Proceedings., pp. 223–230.
16. Nassif, A. B. et al. (2013) "A comparison between decision trees and decision tree forest models for software development effort estimation," in 2013 Third International Conference on Communications and Information Technology (ICCIT), pp. 220–224.
17. Papatheocharous, E. et al. (2017) "An investigation of effort distribution among development phases: A four-stage progressive software cost estimation model," Journal of Software: Evolution and Process. Wiley Online Library, 29(10), p. e1881.
18. Project Management Institute (2013) A guide to the project management body of Knowledge: PMBOK GUIDE. Newtown Square, PA: Project management Institute.
19. Rana, M. E. and Wei, W. (2020) "A Machine Learning based Software Project Schedule Management Solution," Test Engineering and Management, 83, pp. 307–321.
20. Runeson, P. and Höst, M. (2008) "Guidelines for conducting and reporting case study research in software engineering," Empirical Software Engineering, 14(2), p. 131-164
21. Trendowicz, A. and Jeffery, R. (2014) "Classification of Effort Estimation Methods," in Software Project Effort Estimation. Springer, pp. 155–208.
22. Wauters, M. and Vanhoucke, M. (2014) "Support vector machine regression for project control forecasting," Automation in Construction. Elsevier, 47, pp. 92–106.



Κώδικας

Μπορείτε να βρείτε όλο τον κώδικα που δημιουργήθηκε στα πλαίσια της εργασίας στην προσωπική σελίδα μου στο github, ακολουθώντας το παρακάτω link:

<https://github.com/kotzabasis/masterthesis>





Σας ευχαριστώ πολύ για την προσοχή σας!

