HOW EFFECTIVE IS TABU SEARCH TO CONFIGURE SUPPORT VECTOR REGRESSION FOR EFFORT ESTIMATION

A. Corazza, S. Di Martino
University of Naples

F. Ferrucci, C. Gravino, F. Sarro
University of Salerno

E. Mendes
University of Auckland

6th International Conference on Predictive Models in Software Engineering

Outline

- Background and Motivations
  - Software Development Effort Estimation
  - Regression Methods and Support Vector Regression (SVR)
  - SVR for Web Effort Estimation
- Using Tabu Search for configuring SVR
  - Tabu Search
  - The proposed approach
- Case Study
  - Research Goals
  - Dataset Selection, Validation Method and Evaluation Criteria
  - Results
- Conclusions and Future Work

Software Development Effort Estimation

- Software development effort estimation concerns the estimation of the human effort needed to realize a software project
  - effort usually quantified as person-hours or person-months
- Obtaining accurate estimates is a critical activity
  - for planning and monitoring software project development
  - for delivering the product on time and within budget
- Significant over or under-estimates expose a software project to several risks
  - addition of manpower to a late software project makes the project later (Brooks's Law)
  - cancellation of activities, such as documentation and testing, impacts on software quality and maintainability
- Several approaches were proposed to support project managers in estimating software development effort

Regression Methods

- Exploit data from a set of past projects to find the function that best models the training data
  - e.g. linear regression finds the line that minimizes the sum of squared error on training set

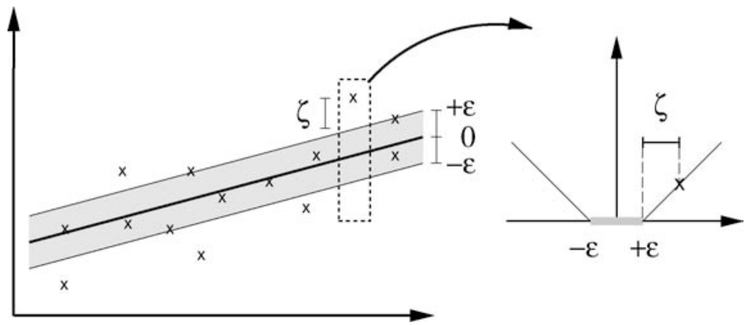
- Such function is then used to estimate the effort for novel projects

- In the effort estimation context data consist of
  - information about some relevant project features (i.e., cost drivers)
  - the effort actually spent to develop the projects

Support Vector Regression (1)

- Support Vector Regression (SVR) is a regression technique based on the Support Vector Machine learning method

- SVR aims to find a function which emulates the traning set points with an error on each point lower than a constant ε
  - among all the possible functions satisfying such constraint the flattest one is chosen
  - a constant C is used to weight errors larger than ε (SVR soft margin version)



Support Vector Regression (2)

- The use of kernel functions allows a better adaption to different problems (linear and non linear)
  - Linear
  - Radial Basis Function (RBF) $K(u,v) = exp(- \gamma\ |u - v|^2\ )$
  - Polynomial $K(u,v) = (s * u \cdot v + c0)degree$
  - …..
- Some parameters need to be specified:
  - Linear Kernel: C and ε
  - RBF: C, ε, and γ
  - Polynomial: C, ε, s, c0,degree

Details about SVR can be found in

B. Scholkopf , Support Vector Learning. R. Oldenbourg Verlag, Munchen. Doktorarbeit, TU Berlin, 1997.
B. Scholkopf, A. Smola, “Learning with Kernels”. 2002, MIT Press
A. J. Smola , B. Schölkopf, “A tutorial on support vector regression”, Statistics and Computing, 14 (3) 2004

Support Vector Regression (3)

- The choice of parameters can have a strong impact on the application of the SVR technique
  - inappropriate setting can lead to over/under-fitting
- No general guidelines are available
  - the appropriate setting could depend on the characteristics of the employed dataset
  - the interaction among parameters complicates the setting
  - the exploration of all the possible settings is not computational affordable as the search space is too large

SVR for Web Effort Estimation (1)

- The use of SVR has been recently proposed for Web Effort Estimation showing promising results
- Corazza *et Al.* case study
  - cross-company database of web projects (Tukutuku)
  - 18 different SVR configurations were exploited
  - best SVR configuration: RBF kernel + log transformation
  - SVR outperformed other widely used techniques
    - Manual Stepwise Regression (MSWR), Case-Based Reasoning (CBR), Bayesian Networks (BN)

A. Corazza, S. Di Martino, F. Ferrucci, C. Gravino, E. Mendes, "Investigating the use of Support Vector Regression for Web Effort Estimation", accepted for publication in Empirical Software Engineering Journal.

SVR for Web Effort Estimation (2)

- The SVR setting can heavily affect the prediction accuracy

- They exploited a semi-automatic approach to set SVR parameters[1]
  - brute-force search with pre-defined steps
  - the information collected during the search were not exploited
  - very expensive computation

[1]Note that in the following we denote it as SVR_{sa}

SVR for Web Effort Estimation (2)

- The SVR setting … rediction accuracy
- They exploite… n to set SVR paramet…
  - brute-force searc… d steps
  - the information c… cted during the search were not exploited
  - very expensive computation

Can we use a suitable heuristic to configure SVR?

[1]Note that in the following we denote it as SVR_{sa}

Outline

- Background and Motivations
  - Software Development Effort Estimation
  - Regression Methods and Support Vector Regression (SVR)
  - SVR for Web Effort Estimation
- Using Tabu Search for configuring SVR
  - Tabu Search
  - The proposed approach
- Case Study
  - Research Goals
  - Dataset Selection, Validation Method and Evaluation Criteria
  - Results
- Conclusions and Future Work

Tabu Search

- Smart optimization strategy based on local search
  - from a current solution *i* a set of neighboring solutions were explored
  - an objective function is used to check which solution is the most suitable to continue the search
  - the search is focused in the most promising area
    - trying to escape from local optima
    - avoiding yet explored area and loops
- Key elements
  - solution structure, moves, objective function, tabu list, aspiration criteria, termination criteria

The proposed approach (1)

- **Solution structure:** an SVR parameter setting
  - Linear Kernel: S =(C, ε) , C $\in$ (0,1000], $\varepsilon \in$ (0,0.1]
  - RBF Kernel: S =(C, ε , γ) , $\gamma \in$ (0,0.1]
- The search starts from a random solution
  - at each iteration TS applies local transformations (moves) to the current solution *S* defining 25 neighboring solutions
- **Move**: random variation of the current solution *S*
  - each parameter of S is modified with a probability of 0.5
  - new parameters are calculated by applying an arithmetic operator {+,−, *, /} to the old parameters and a random number *r*

The proposed approach (2)

- At each step neighboring solutions are compared
  - executing SVR with the corresponding settings
  - using an objective function to evaluate the accuracy of the obtained estimates

- **Objective function**: mean of MMRE and MEMRE
  - MMRE: Mean of Magnitude of Relative Error (MRE), where

$$MRE = \frac{|ActualEffort - EstimatedEffort|}{ActualEffort}$$

  - MEMRE: Mean of Estimated MRE (EMRE), where

$$EMRE = \frac{|ActualEffort - EstimatedEffort|}{EstimatedEffort}$$

The proposed approach (3)

- **Tabu List:** record recent moves marking them as taboo
  - the use of a taboo move is forbidden for seven iterations
- **Aspiration criteria:** a taboo move is allowed only if it results in a solution with an objective value better than the one of the current best-known solution
- **Termination criteria:** the search is stopped after a fixed number of iterations (i.e., 100) were performed

Outline

- Background and Motivations
  - Software development effort estimation
  - Regression Methods and Support Vector Regression (SVR)
  - SVR for Web Effort Estimation
- Using Tabu Search for configuring SVR
  - Tabu Search
  - The proposed approach
- Case Study
  - Research Goals
  - Dataset Selection, Validation Method and Evaluation Criteria
  - Results
- Conclusions and Future Work

Case Study: Research Goals

- RG_1: Is Tabu Search able to effectively set SVR configuration parameters?
  - comparison with SVRsa
- RG_2: Are the effort predictions obtained using the combination of TS and SVR significantly superior to the ones obtained by other techniques?
  - comparison with widely used estimation methods

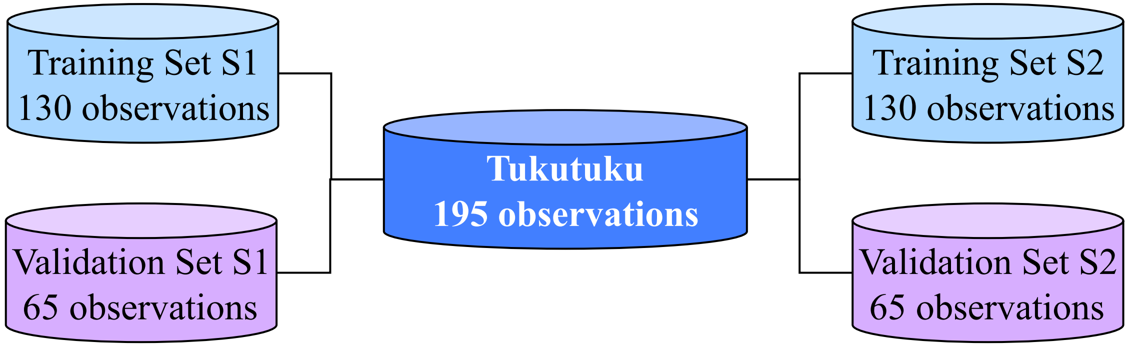
Case Study: Dataset Selection

- Tukutuku dataset: data on 195 industrial Web Projects from 10 countries
  - Data mainly (85%) collected in controlled way
  - Project Types
    - new developments (65.6%) and enhancement projects (34.4%)
    - some old web sites, many new applications
  - Many technologies
    - PHP 42.6% of the projects, ASP VBScript or .Net 13.8%, Perl 11.8%, J2EE 9.2%, other solutions 9.2%, the remaining projects used only HTML and/or Javascript
  - Project size ranges from very small to very large

Case Study: Validation Method

- Hold-Out validation
  - we randomly split the original dataset into
    - a training set for model building
    - a validation set for model evaluation
- 2 random splits used in previous researches

A. Corazza, S. Di Martino, F. Ferrucci, C. Gravino, E. Mendes, "Investigating the use of Support Vector Regression for Web Effort Estimation", accepted for publication in Empirical Software Engineering Journal.



Case Study: Evaluation Criteria

- Accuracy measures
  - MMRE (Mean of MRE)
  - MdMRE (Median of MRE)
  - Pred(25) (Prediction at level 25)
  - MEMRE (Mean of the Estimated MRE)
  - MdEMRE (Median of the Estimated MRE)
- Boxplots of absolute residuals
  - Summarize the data through a visual representation
  - $\text{Absolute residuals} = \left| \text{ActualEffort} - \text{EstimatedEffort} \right|$
- Statistical significance test of absolute residuals
  - Wilcoxon Test (α = 0.05) since the absolute residuals were not normally distribuited and the data was naturally paired

Case Study: SVR Configurations

- Two kernels were used: Linear and RBF
- Two strategies for input data
  - no preprocessing
  - pre-processing*: logarithmic transformation of features
- Thus, 4 SVR configurations were considered:
  - C1 Lin: Linear Kernel+ No features transformation
  - C1 RBF: RBF Kernel + No features transformation
  - C3 Lin: Linear Kernel + Log features transformation
  - C3 RBF: RBF Kernel + Log features transformation

* to avoid that large differences in the values of the features could have the unwanted effect of giving greater importance to some of the examples with respect to the others. This is especially true for cross-company datasets

Case Study: TS+SVR vs. Other Estimation Methods

- Manual Stepwise Regression: MSWR
- Case Based Reasoning (One, Two, and Three Analogies): CBR1, CBR2, and CBR3
- Bayesian Networks (Automatically generated model, Hybrid model): BNAuHu, BNHyHu
- Mean Effort
- Median Effort

A. Corazza, S. Di Martino, F. Ferrucci, C. Gravino, E. Mendes, "Investigating the use of Support Vector Regression for Web Effort Estimation", accepted for publication in Empirical Software Engineering Journal.

Case Study: Results
TS+SVR vs. SVR_{sa}

First validation set – Summary Measures TS + SVR

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	1.989	0.582	0.234	1.357	0.732
C1 RBF	1.712	0.663	0.297	1.397	0.735
C3 Lin	1.151	0.552	0.281	1.122	0.530
C3 RBF	0.590	0.339	0.391	0.689	0.365

First validation set – Summary Measures SVR_{sa}

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	1.984	0.583	0.234	1.355	0.735
C1 RBF	1.422	0.777	0.203	2.979	0.733
C3 Lin	1.151	0.555	0.281	1.118	0.555
C3 RBF	0.591	0.411	0.344	0.603	0.467

Second validation set – Summary Measures TS + SVR

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	2.443	0.700	0.231	2.209	0.716
C1 RBF	2.733	0.954	0.200	2.855	0.692
C3 Lin	1.007	0.539	0.231	0.764	0.491
C3 RBF	0.856	0.455	0.400	0.498	0.410

Second validation set – Summary Measures SVR_{sa}

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	2.298	0.826	0.077	2.137	0.754
C1 RBF	2.723	0.751	0.108	7.804	0.886
C3 Lin	1.215	0.450	0.338	0.680	0.542
C3 RBF	0.910	0.360	0.415	0.506	0.413

A. Corazza, S. Di Martino, F. Ferrucci, C. Gravino, E. Mendes, "Investigating the use of Support Vector Regression for Web Effort Estimation", accepted for publication in Empirical Software Engineering Journal.

Case Study: Results
TS+SVR vs. SVR_{sa}

First validation set – Summary Measures TS + SVR

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	1.989	0.582	0.234	1.357	0.732
C1 RBF	1.712	0.663	0.297	1.397	0.735
C3 Lin	1.151	0.552	0.281	1.122	0.530
C3 RBF	0.590	0.339	0.391	0.689	0.365

First validation set – Summary Measures SVR_{sa}

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	1.984	0.583	0.234	1.355	0.735
C1 RBF	1.422	0.777	0.203	2.979	0.733
C3 Lin	1.151	0.555	0.281	1.118	0.555
C3 RBF	0.591	0.411	0.344	0.603	0.467

Second validation set – Summary Measures TS + SVR

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	2.443	0.700	0.231	2.209	0.716
C1 RBF	2.733	0.954	0.200	2.855	0.692
C3 Lin	1.007	0.539	0.231	0.764	0.491
C3 RBF	0.856	0.455	0.400	0.498	0.410

Second validation set – Summary Measures SVR_{sa}

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	2.298	0.826	0.077	2.137	0.754
C1 RBF	2.723	0.751	0.108	7.804	0.886
C3 Lin	1.215	0.450	0.338	0.680	0.542
C3 RBF	0.910	0.360	0.415	0.506	0.413

- VS1: Comparable results when Linear Kernel was used
- TS+SVR achieved better results for both validation sets when RBF Kernel was used
- The Wilcoxon Tests revealed that the abs residuals of C3 RBF (VS1, VS2) and C1 Lin (VS2) obtained with TS+SVR were significantly less than those of SVR_{sa}

Case Study: Results
SVR Configurations

First validation set – Summary Measures

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	1.989	0.582	0.234	1.357	0.732
C1 RBF	1.712	0.663	0.297	1.397	0.735
C3 Lin	1.151	0.552	0.281	1.122	0.530
C3 RBF	**0.590**	**0.339**	**0.391**	**0.689**	**0.365**

First validation set – Wilcoxon Tests

	C1 Lin	C1 RBF	C3 Lin	C3 RBF
C1 Lin	-	No(0.055)	No (0.836)	-
C1 RBF	-	-	-	-
C3 Lin	-	**Yes(0.031)**	-	-
C3 RBF	**Yes(0.014)**	**Yes(0.000)**	**Yes(0.001)**	-

Second validation set – Summary Measures

	MMRE	MdMRE	Pred(25)	MEMRE	MdEMRE
C1 Lin	2.443	0.700	0.231	2.209	0.716
C1 RBF	2.733	0.954	0.200	2.855	0.692
C3 Lin	1.007	0.539	0.231	0.764	0.491
C3 RBF	**0.856**	**0.455**	**0.400**	**0.498**	**0.410**

Second validation set – Wilcoxon Tests

	C1 Lin	C1 RBF	C3 Lin	C3 RBF
C1 Lin	-	No(0.253)	No (0.925)	-
C1 RBF	-	-	-	-
C3 Lin	-	-	-	-
C3 RBF	**Yes(0.000)**	**Yes(0.001)**	**Yes(0.007)**	-

- C3 RBF is characterized by better summary values for both validation sets
- The Wilcoxon Test revelaed that the absolute residuals of C3 RBF are significantly less than those of the other configurations for both validation sets

Case Study: Results
TS+SVR vs. other estimation methods (1)

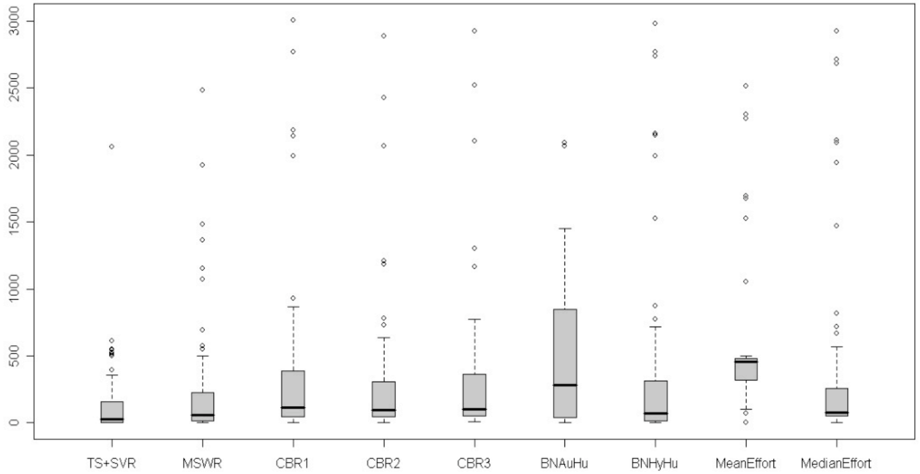
SUMMARY MEASURES ACHIEVED USING DIFFERENT METHODS

	First validation set					Second validation set				
	MMRE	**MdMRE**	**Pred(25)**	**MEMRE**	**MdEMRE**	**MMRE**	**MdMRE**	**Pred(25)**	**MEMRE**	**MdEMRE**
TS+SVR (C3 RBF)	**0.59**	**0.34**	**0.39**	**0.69**	**0.36**	0.86	**0.46**	**0.40**	**0.50**	**0.41**
MSWR	1.50	0.64	0.23	1.36	0.64	**0.73**	0.66	0.11	2.86	1.21
CBR1	5.27	0.97	0.08	31.70	3.43	4.46	0.92	0.08	21.81	0.95
CBR2	5.06	0.87	0.11	3.59	0.81	6.73	0.89	0.15	15.65	0.90
CBR3	5.63	0.97	0.09	4.17	0.88	6.09	0.84	0.09	13.26	0.89
BNAuHu	7.65	1.67	7.69	1.07	0.76	4.09	0.96	0.02	7.90	0.93
BNHyHu	1.90	0.86	0.15	13.06	2.38	27.95	5.31	0.09	1.34	0.90
MedianEffort	5.02	0.93	0.09	4.43	0.94	4.95	0.89	0.15	4.62	0.78
MeanEffort	30.35	3.99	0.08	1.07	0.91	27.94	5.31	0.03	1.34	0.90

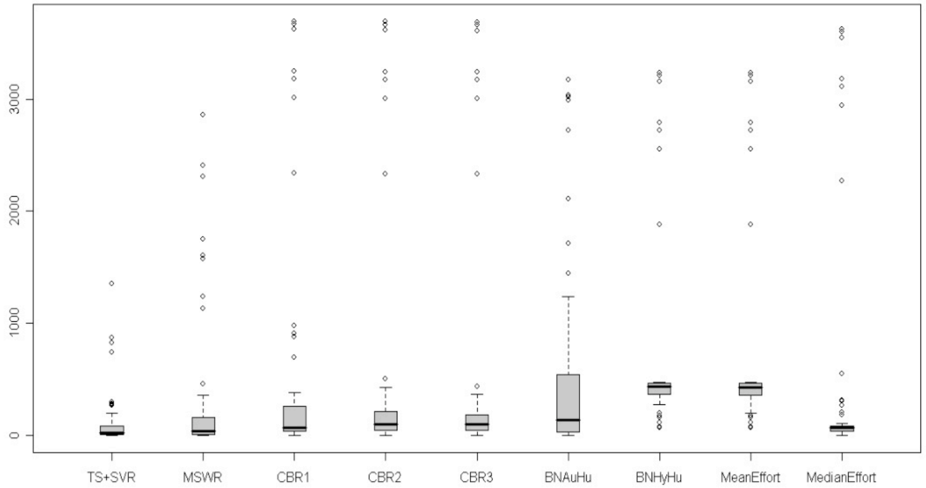
- TS+SVR is characterized by better summary measures respect to all the other techniques

Case Study: Results
TS+SVR vs. other estimation methods (2)

- TS+SVR boxplots were characterized by
  - box lenght and tails less skewed than others
  - median closer to zero respect to the others
  - outliers closer to the tails than the outliers of other methods



Boxplots of abs residuals for the 1st validation set



Boxplots of abs residuals for the 2nd validation set

Case Study: Results
TS+SVR vs. other estimation methods (3)

RESULTS OF THE WILCOXON TEST COMPARING DIFFERENT METHODS – First validation set

	MSWR	CBR1	CBR2	CBR3	BNAuHu	BNHyHu	Mean Effort	Median Effort
TS+SVR (C3 RBF)	**Yes (0.044)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.025)**	**Yes (0.000)**	**Yes (0.000)**
MSWR	-	**Yes (0.000)**	**Yes (0.002)**	**Yes(0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**
CBR1		-	**Yes (0.052)**	No (0.398)	No (0.288)		**Yes (0.022)**	
CBR2			-	No (0.227)	**Yes (0.022)**	No (0.229)	**Yes (0.000)**	No (0.335)
CBR3				-	**Yes (0.023)**	No (0.288)	**Yes (0.000)**	No (0.422)
BNAuHu					-		No (0.223)	No (0.822)
BNHyHu		**Yes (0.000)**			**Yes (0.038)**	-	**Yes (0.022)**	
Mean Effort							-	
Median Effort		**Yes (0.003)**				**Yes (0.042)**	**Yes (0.002)**	-

RESULTS OF THE WILCOXON TEST COMPARING DIFFERENT METHODS – Second validation set

	MSWR	CBR1	CBR2	CBR3	BNAuHu	BNHyHu	Mean Effort	Median Effort
TS+SVR (C3 RBF)	**Yes (0.046)**	**Yes (0.000)**	**Yes (0.001)**	**Yes (0.001)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.010)**
MSWR	-	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.001)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.009)**
CBR1		-	No (0.909)	No (0.893)	No (0.241)	**Yes (0.009)**	**Yes (0.009)**	
CBR2			-		No (0.380)	**Yes (0.000)**	**Yes (0.000)**	
CBR3			**Yes (0.000)**	-	No (0.264)	**Yes (0.000)**	**Yes (0.000)**	
BNAuHu					-	**Yes (0.000)**	**Yes (0.000)**	
BNHyHu						-	**Yes (0.000)**	
Mean Effort							-	
Median Effort		**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	**Yes (0.000)**	-

Outline

- Background and Motivations
  - Software development effort estimation
  - Regression Methods and Support Vector Regression (SVR)
  - SVR for Web Effort Estimation
- Using Tabu Search for configuring SVR
  - Tabu Search
  - The proposed approach
- Case Study
  - Research Goals
  - Dataset Selection, Validation Method and Evaluation Criteria
  - Results
- Conclusions and Future Work

Conclusions and Future Work

- TS is effective for configuring SVR in estimating Web application development effort
  - automatic choice of the parameters required to run SRV
  - significant improvement on prediction accuracy respect to SVR_{sa} and other estimation techniques
- The results encourage for further investigation
  - of the proposed approach on other datasets
  - combinations with other estimation techniques

Questions?

Thanks for your attention

Federica Sarro
fsarro@unisa.it

Department of Mathematics and Informatics

University of Salerno