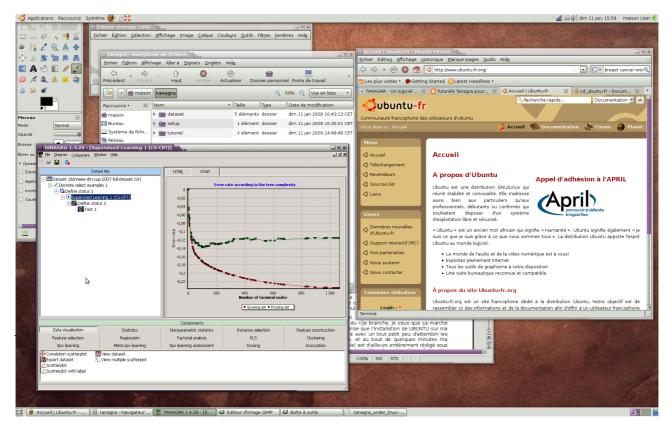
Introduction

How to install and run Tanagra under Linux.

The users ask sometimes "Can I use Tanagra under Linux?" The answer is YES and NO.

NO, we cannot execute natively Tanagra under Linux. It is a 32-bits program for Windows.

But YES, we can run Tanagra under Linux using WINE, a famous Linux application which allows us to run Windows programs on Linux. We can then take all the advantages of Tanagra without asking any questions about compatibilities.



In this tutorial, we show how to install and use Tanagra under Ubuntu (<u>http://www.ubuntu.com/</u>, a free of charge version of Linux) by the way of WINE (<u>https://help.ubuntu.com/community/Wine</u>) that we must install before.

All the steps described in this tutorial are tested under the latest French version of Ubuntu (8.10). I expect that there are not many differences for the others distributions. I am not really an expert in this domain.

Installing WINE

Ubuntu must be operational on your computer. To install Wine, we follow the descriptions on this website \rightarrow <u>https://help.ubuntu.com/community/Wine</u>

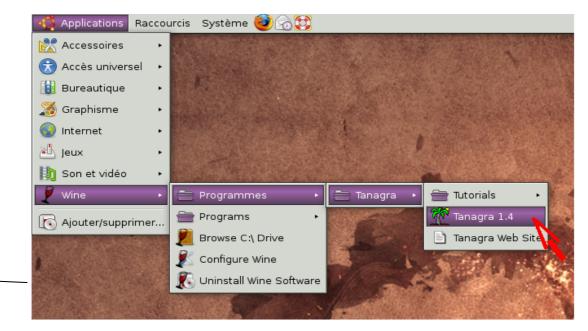
Tanagra

Installing Tanagra

When Wine is properly configured, we download the latest version of the SETUP of Tanagra on the website (http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html, see the Download section). This link (https://help.ubuntu.com/community/Wine#Installing%20Windows%20Applications%20Using%20Wine) describes the main steps of the installation using the "setup_tanagra.exe" program. We select the default "c:\program files\tanagra" directory.

Running Tanagra using Wine

After we restart Ubuntu, Tanagra is now available in the main menu of the Desktop. For the French version, the location is APPLICATIONS / WINE / PROGRAMMES. To launch Tanagra, we click on the icon.



Tanagra is normally started. We can use it without special precautions. All features are operational.

We can import and use TEXT file (tab separator), WEKA file (ARFF) or EXCEL file (XLS, it is not necessary that the spreadsheet should be installed on your computer, see http://data-miningtutorials.blogspot.com/2008/10/excel-file-format-direct-importation.html).

Tanagra					
Approx/WF Pastourum Sysème Coll Accessures Image: Im	Si Decri Sino HA Di Si Decrificacean	atmpag			
	Liato estelaction	حاظها	Components Nor carametric statistics	Instance spection	Feature construction
	Headh (e se aichte f Obail ean àng	Kegnessimn Meta-spyllearn ng	Factorial analysis Storiganistic assessment :	- S Elucing	Clinite ing Association
	 Ione at on scatterplot Ione at on scatte	k paceed Kimultiple spatterpipt			

In the following screen shot, we show the utilization of the A PRIORI PT component. This example is symbolic because we call an external program in this situation (BORGELT's "apriori.exe", <u>http://www.borgelt.net/apriori.html</u>). The data transfer and the recovery of the results operate properly.

							THE	oriels Tanagra pour le Data Mining: Règles d'association avec les F	REFIX TREE - Mozilla Firefox		
							Echier Édition Affichage Historique Marque-pages Qutils Aide				
							🗘 🗘 🔹 🛞 🖉 🛐 🖸 http://tutoriels-data-mining.blogspot.com/2008/04/rgles-dass(💦 🗋 💽 💽 capture of				
								🔚 🛅 Les plus visités 🔹 🏚 Getting Started 🔊 Latest Headlines 🔹			
ANAGRA 1.4.29 - [A priori PT 1]							💶 🗶 🗚 AGRA - Un logiciel 💥 🕒 Tutoriels Tanagra pour 💥 🧔 wine - Documentation 💥 🗔 cd_ubuntu-fr - Docu				
e Diagram Component Window	Help						_ 코×	[RECHERCHER LE BLOG] MARQUER LE	BLOG Blog suivant» Créer un blog Co		
Default title			ecution log								
Dataset (banque:txt)	C1Program Files[Tanagaleevelpatoiteae - find association rules with the patoin algorithm version 4.31 (2007.03.12) (C) (D) 96-007 Orthites Borget reading C1 windows[temp[dat6561.thrp[25] fisem(s), 198 transaction(s)] done [0.01s]. filtering, scritting and reading fismes[13 fisem(s)] done [0.00s]. reating transaction thesdite [0.00s]. checking absolute of rait 0.2.3 done [0.00s]. writing 2:\home[matorittingan]dataset(output.rul[54 rule(s)] done [0.00s].							Tuttoriels Tanagra pour le Data Mining Ce biogreense les diductivés pour Tanagra, lis sont organicés en calégories. On most cela: Chaque andre est accompagné duri totte de présentation, dure les de de vers la diductivé (off) de de à solvique de Dana commer ce (calégorie : Tanagra avec d'aures legiciés titres prime, Orange, R, Rapublinet, Spins, Vesa) or comm	nots-clés, du lien vers les données, du lien et les autres »), nous montrons comment faire		
		Rules (#54	association rules loaded]								
	N° Antecedent		Consequent		t Confi	Lift 🔻		MERCREDI 2 AVRIL 2008	Supports et tutoriels		
	54 port_action=oui∧csp 37 port_action=oui∧hat	i≈cadre_moyen ∧ habit+locataire it+locataire	accord+oui accord=oui	34,3 40,4	86,1 86,0	119,2 119,1	-	Règles d'association avec les PREFIX TREE	🕱 Page principale du blog		
	12 port_action=oui		accord=oui	44,9	83,2	115,2		La construction de règles d'association pose de problèmes de performances,	🕷 Cours Data Mining		
	35 port_action+oui ∧ csp	=cadre_moyen	accord=oui	37,4	83,1	115,1		tant en occupation mémoire qu'en temps de traitement. L'implémentation	🗷 Portail Data Mining		
	2 Age=ancien 25 Age=ancien ∧ habit=l	and size	csp=cadre_moyen csp=cadre_moyen	35,9 34,8	95,9 95,8	111,1		actuelle dans TANAGRA est relativement rapide, en revanche, elle est très	😹 Tutoriels en anglais		
	45 sitiam+marie ∧ csp=c		accord=oui	35,4	77,8	107,7		gourmande en mémoire, au point de la saturer très rapidement dès que l'on a à produire un grand nombre de règles. De plus, l'affichage des règles en HTML			
	47 sittam=marie ∧ habit=	locataire	accord+oui	38,4	77,6	107,4		est tributaire du composant d'affichage, un peu limité, au point que le temps	Logiciels		
	8 revenu=tranche_2 28 port_action=non 八 cs		csp=cadre_moyen habit=locataire	43,9 40,4	92,6 97,6	107,2		consacré à l'affichage est parfois aussi important que le temps consacré à	🗷 Site du logiciel Tanagra		
	31 revenu=tranche_2 A		csp=cadre_moyen	40,4	92,0	106,6		l'élaboration des règles.	🕱 Téléchargement Tanagra		
	3 Age=ancien		habit=locataire	36,4	97,3	106,4		Il fallait donc se tourner vers un module très performant de construction des	🕷 Site du logiciel Sipina		
	24 Age=ancien A csp=c 7 pot_action=non	adre_moyen	habit=locataire habit=locataire	34,8 44,4	97,2 96,7	106,3		règles et proposer une nouvelle fenêtre d'affichage peu sensible au nombre de			
		in ∧ csp=cadre_moyen	habit=locataire	40.9	96.4	105.5		règles, fussent-elles de plusieurs centaines de milliers.			
	29 port_action=non 八 ha	bit-locataire	csp=cadre_moyen	40,4	90,9	105,3			Catégories des tutoriels		
	11 demand=consommatic 17 sittam=marie	n	habit=locataire accord=oui	47,0 42,4	95,9 75,7	104,9		Sur la création des règles, j'ai découvert les travaux de Christian BORGELT sur les PREFIX TREE. Il propose une implémentation réellement impressionnante.	Analyse discriminante (10)		
				34,3		104,8		Traduire le code en DELPHI m'exposait aux risques de mauvaises	Analyse factorielle (11)		
	6 port_action=non		csp=cadre_moyen	41,4		104,3		interprétations de son travail, et donc d'introduction d'erreurs ; le passage par	App. Supervisé - Scoring (35)		
	5 Age=mature 43 enfant+zero // habit+i	nostaine.	habit=locataire csp=cadre_moyen	39,4 43,4	95,1 88,7	104,1		des DLL est également séduisant mais m'oblige à faire un travail de traduction des structures en C vers DELPHI, toujours hasardeux, pour la définition des	Arbres de décision (16)		
	51 accord=oui // habit=lo		csp=cadre_moyen	59,1	88,6	102,6		unités d'import. J'ai donc décidé d'inaugurer une nouvelle approche avec cette	Classification - Clustering (12)		
	50 accord=oui /\ csp=ca	die_moyen	habit=locataire	59,1	93,6	102,4		nouvelle version, l'appel à un programme externe avec passage de fichiers			
	9 revenu=tranche_2 40 enfant+zero /\ accord	kari	habit=locataire csp=cadre_moyen	44,4 33.3	93,6 88,0	102,4		temporaires. La rapidité de l'ensemble dépend en grande partie du temps	Construction de variables (5)		
	30 revenu=tranche_2 A		habit=locataire	40,9		101,9	-	consacré à l'écriture et à la lecture des fichiers temporaires. Force est de constater que le travail de BORGELT est réellement impressionnant. Au final,	Importation des données (8)		
	-10-	Components						cette approche semble viable, du moins tant qu'il n'est pas nécessaire de	Règles d'association (6)		
Data visualization	Statistics	Nonparametric statistics	Instance selection		Fea	ture construc	tion	produire des données qui seront par la suite utilisées dans le diagramme. Nous	Régression (7)		
Feature selection	Regression	Factorial analysis	PLS		Clustering			en montrons un exemple dans ce didacticiel.	Régression logistique (7)		
Spv learning	Meta-spy learning	Spv learning assessment	Scoring	Γ		Association		L'autre point important était de créer une fenêtre de visualisation des règles qui	Régression PLS (6)		
priori 🕄 Spv Assoc								ne s'effondre pas dès que leur nombre excède la centaine de milliers de règles,	Sipina (11)		
priori MR E Spy Assoc	Tree							et qui par ailleurs, comporte des fonctionnalités de tri selon différents critères.	Statistiques et tests (23)		
soc Outlier								Nous avons donc élaboré un outil simple qui permet de récupérer les sorties de BORGELT et d'afficher simplement les règles dans une fenètre convivale.	Sélection de variables (5)		
				•				berteter et a anenar ampientent na regies dans une leitede commane.	Tanagra - Développement (2)		
							3	é			
10-00 CO. 10-00-00-00-00-00-00-00-00-00-00-00-00-0			State College And	1000	1000	Sec. Sec.	STATISTICS.				

13/01/2009

Conclusion

Tanagra is an application built with DELPHI for WIN32. One might think that its use is therefore confined exclusively to Windows. We see in this tutorial that with the installation of WINE, a well known tool of the Linux users, we can fully use Tanagra in the Linux environment.