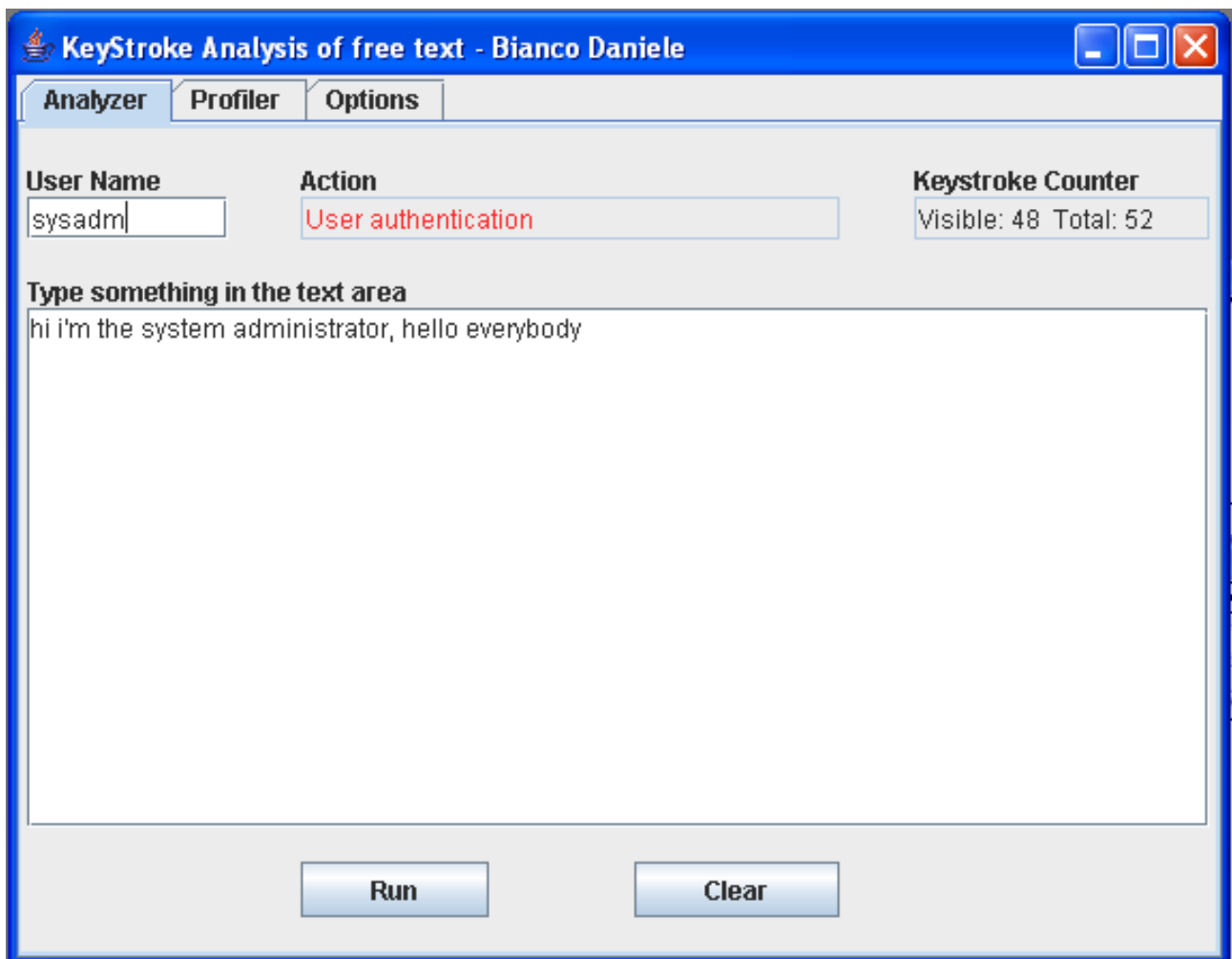


KEYSTROKE ANALYSIS OF FREE TEXT

Questo applicativo implementa una tecnica di analisi dei ritmi di battitura degli utenti, usati come misura biometrica per il riconoscimento di identità. Il software permette di raccogliere campioni di testo sia da aggiungere a un profilo utente già esistente sia da confrontare con quelli precedentemente memorizzati, con la possibilità di specificare i parametri del calcolo della misura delle distanze. L'applicativo fornisce un ambiente di lavoro grafico in cui l'utente può svolgere le operazioni con il mouse e immettere testo tramite tastiera. I dati raccolti nel campionamento vengono registrati dentro la directory dell'utente in questione, e vengono salvati, per ogni campione, sia l'originale sequenza codificata dei tasti premuti e il loro timing, sia gli n-grafi, suddivisi in file diversi a seconda della dimensione di n. Ecco di seguito le modalità d'uso e le funzionalità del software.

Il pannello Analyzer



Questo pannello permette di compiere le operazioni di classificazione e autenticazione di un utente attraverso la composizione di un testo libero. L'interfaccia si presenta così:

Campo User-Name

Se lasciato vuoto, il sistema provvedera', coerentemente con il valore di k scelto nel pannello delle opzioni, a classificare o identificare il campione; se il nome inserito invece rappresenta un utente che ha completato la fase di training, e $k \neq 1$, il sistema provvedera' all' autenticazione dello specifico utente.

Campo Action

Qui vengono visualizzate informazioni utili all' utilizzatore, nello specifico l' azione successiva che il sistema eseguirà sulla base dei dati inseriti.

Campo Keystroke Counter

Mostra all' utente sia quanti caratteri visibili sono stati inseriti nella input area, sia quelli effettivi memorizzati. Ricordiamo che i due valori possono essere diversi a causa della correzione di eventuali errori di inserimento effettuata tramite il tasto backspace.

Campo Input Area

Qui l' utente deve inserire una porzione di testo: può liberamente scrivere quello che gli pare, nel modo più naturale possibile, un testo di fantasia oppure ricopiare un testo esistente: l' importante è che abbia un senso compiuto nella lingua in cui è stato scritto. È libero di fare errori tipografici, di correggerli, oppure no. Il testo affinché sia utile dovrebbe essere lungo circa 700-800 caratteri effettivi. Naturalmente si possono fare delle pause, ad esempio per rileggere quanto è stato scritto.

Pulsante Run

Viene eseguita l' azione specificata nel campo action.

Pulsante Clear

La schermata viene resettata.

Il pannello Profiler

The screenshot shows a software window titled "KeyStroke Analysis of free text - Bianco Daniele". It has three tabs: "Analyzer", "Profiler", and "Options", with "Profiler" selected. The interface includes three input fields at the top: "User Name" containing "dummy", "Action" containing "It will be created a new user", and "Keystroke Counter" containing "Visible: 63 Total: 63". Below these is a large text area with the text "hi i'm doing the profiling phase.. i'm a new user in the system". At the bottom, there are two buttons: "Add the sample" and "Clear".

Questo pannello permette di inserire un nuovo utente, oppure di aggiungere un campione di testo a un utente già registrato; serve quindi a registrare il profilo degli utenti del sistema, e a completare il training necessario affinché il software possa operare con maggiore affidabilità.

Campo user-name

Qui deve essere inserito il nome di un utente, nuovo oppure già esistente nel database; nel primo caso, verrà visualizzata una finestra per confermare la creazione di un nuovo utente.

Campo Action

Viene visualizzata l'operazione che il sistema eseguirà quando verrà premuto Add this sample; nello specifico, indicherà "creazione di un nuovo utente" oppure "aggiunta di un campione per questo utente".

Campo Keystroke Counter

Stessa funzionalità dell'omonimo campo nel pannello Analyzer.

Campo input area

Anche qui l'utente deve inserire una porzione di testo: valgono le stesse considerazioni fatte per il pannello Analyzer.

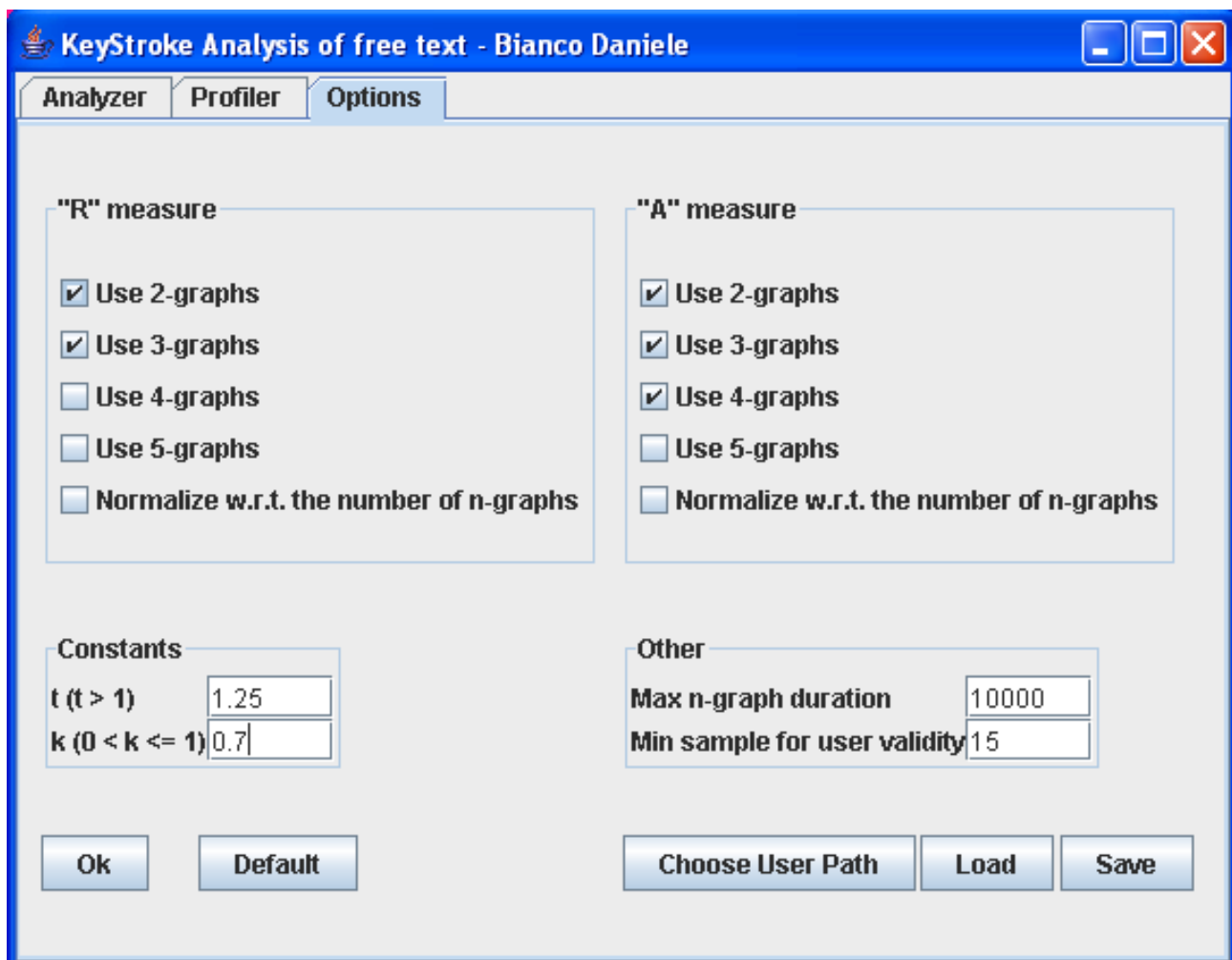
Pulsante Add the Sample

Viene aggiunto un campione all'utente in questione; se non esiste nessun utente con il nome specificato, viene creata una nuova directory con il nuovo campione.

Pulsante Clear

La schermata viene resettata.

Il pannello Options



Questo pannello permette di specificare le opzioni che il sistema deve utilizzare, ad esempio nel calcolo delle misure R ed A, ma non solo:

Campi R – Measure

Qui si possono specificare quali tipi di n-grafi debbano essere presi in considerazione

dal sistema, e in che modo pesare i valori della misurazione. Come si puo' vedere, le combinazioni possibile sono moltissime, e i valori risultanti possono essere molto diversi: utenti autenticati usando tutti e quattro i tipi di n-grafi possono venire rifiutati dal sistema se si usano settaggi diversi, e viceversa. Si rimanda al capitolo 4 e al documento 1 della bibliografia per maggiori informazioni sui risultati ottenuti utilizzando diverse misure.

Campo Constants

In questa area e' possibile specificare le costanti che il sistema deve usare nel calcolo di R ed A. Il sistema non accetta valori di $T \leq 1$, e valori di K non compresi nell'intervallo (0,1].

Campo Other

In questa area si possono specificare le seguenti opzioni:

- *max n-graph duration*: indica la midurata massima possibile di un n-grafo, oltre il quale viene ignorato.

- *min sample for user validity*: indica il numero minimo di campioni che un utente deve inserire affinche' il training sia completo

Pulsante Default

Ripristina i settaggi di default del sistema, oppure, se presente, quelli memorizzati nel file default.opt nella directory dell' applicativo.

Pulsante Load

Permette di caricare le opzioni da un file precedentemente salvato.

Pulsante Save

Permette di salvare le opzioni correnti in un file. Notare che se come nome si specifica "default.opt" le opzioni verranno caricate di default al successivo riavvio dell'applicativo.

Pulsante Choose User Path

Permette di scegliere la cartella che contiene i profili utenti del sistema. Di default e' la subdirectory "users".

Modalita' Operative

E' possibile lanciare l' applicativo in due modalita': direttamente, oppure dal prompt dei comandi; nel primo caso, il sistema provvedera' a fornire il risultato dell' operazione richiesta tramite delle finestre di dialogo che permetteranno di salvare il campione di testo, e similmente segnalera' eventuali errori, ad esempio valori invalidi per le costanti t e k.

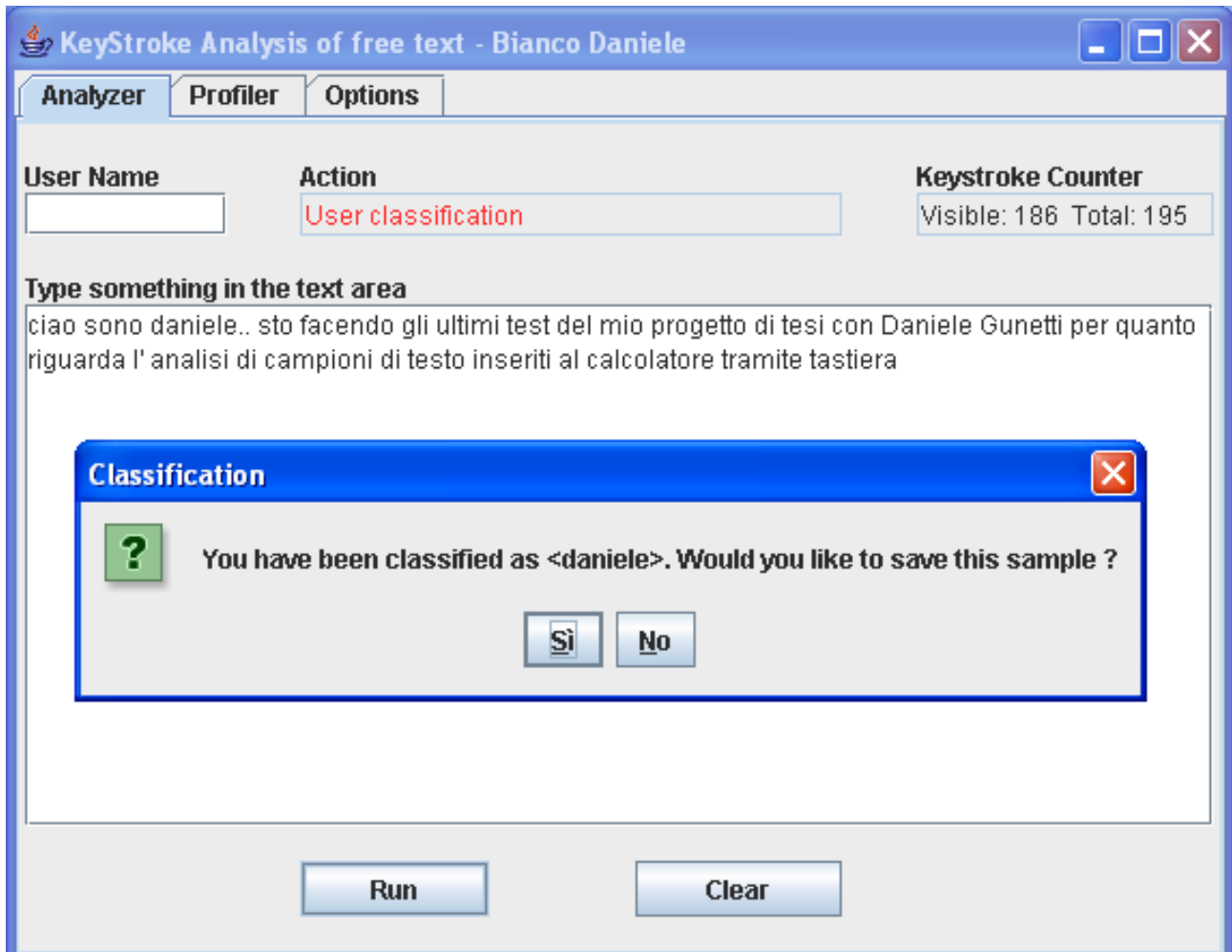
Nel secondo caso, verranno anche visualizzate a schermo delle stampe che mostrano i risultati delle operazioni compiute dal sistema, tra cui: calcolo della misura di R ed A con l' indicazione, per ogni n e per ogni utente, degli n-grafi condivisi e del valore della misura; i risultati delle operazioni di classificazione, autenticazione e identificazione, con, dove opportuno, indicazione della misura della distanza media per ogni utente dal campione inserito, misura delle distanze tra i profili memorizzati per un dato utente, e nome e distanza media per il secondo utente piu' vicino al campione inserito.

```
C:\WINDOWS\system32\cmd.exe - java KeyStrokeAnalyzer
R2= 0.433268 (sh 78); R3= 0.370064 (sh 66); R4= 0.354762 (sh 41); R5= 0.500000 (sh 25);
A2= 0.615385 (sh 78); A3= 0.500000 (sh 66); A4= 0.341463 (sh 41); A5= 0.440000 (sh 25);
SYSTEM: Comparison with sample <gatti-2> :
R2= 0.560897 (sh 79); R3= 0.566327 (sh 56); R4= 0.532967 (sh 27); R5= 0.533333 (sh 11);
A2= 0.645570 (sh 79); A3= 0.500000 (sh 56); A4= 0.555556 (sh 27); A5= 0.545455 (sh 11);
SYSTEM: Comparison with sample <gatti-3> :
R2= 0.445102 (sh 78); R3= 0.489712 (sh 54); R4= 0.695833 (sh 31); R5= 0.569444 (sh 17);
A2= 0.615385 (sh 78); A3= 0.555556 (sh 54); A4= 0.580645 (sh 31); A5= 0.647059 (sh 17);
SYSTEM: Comparison with sample <gatti-4> :
R2= 0.439103 (sh 79); R3= 0.588970 (sh 62); R4= 0.663743 (sh 37); R5= 0.785124 (sh 22);
A2= 0.607595 (sh 79); A3= 0.612903 (sh 62); A4= 0.594595 (sh 37); A5= 0.636364 (sh 22);
SYSTEM: Comparison with sample <gatti-5> :
R2= 0.454791 (sh 77); R3= 0.518462 (sh 51); R4= 0.597633 (sh 26); R5= 0.755102 (sh 14);
A2= 0.571429 (sh 77); A3= 0.529412 (sh 51); A4= 0.653846 (sh 26); A5= 0.785714 (sh 14);
SYSTEM: Comparison with sample <gatti-6> :
R2= 0.537179 (sh 79); R3= 0.539855 (sh 47); R4= 0.537190 (sh 22); R5= 0.640000 (sh 10);
A2= 0.632911 (sh 79); A3= 0.468085 (sh 47); A4= 0.318182 (sh 22); A5= 0.900000 (sh 10);
SYSTEM: Comparison with sample <gatti-7> :
R2= 0.489204 (sh 77); R3= 0.454545 (sh 45); R4= 0.432099 (sh 18); R5= 1.000000 (sh 5);
A2= 0.675325 (sh 77); A3= 0.555556 (sh 45); A4= 0.555556 (sh 18); A5= 1.000000 (sh 5);
SYSTEM: Comparison with sample <gatti-8> :
R2= 0.492308 (sh 79); R3= 0.582500 (sh 40); R4= 0.506173 (sh 18); R5= 0.833333 (sh 5);
A2= 0.569620 (sh 79); A3= 0.600000 (sh 40); A4= 0.555556 (sh 18); A5= 0.800000 (sh 5);
SYSTEM: Comparison with sample <gatti-9> :
R2= 0.446417 (sh 78); R3= 0.524235 (sh 56); R4= 0.516484 (sh 27); R5= 0.642857 (sh 15);
A2= 0.538462 (sh 78); A3= 0.589286 (sh 56); A4= 0.629630 (sh 27); A5= 0.800000 (sh 15);
SYSTEM: Comparison with sample <gatti-10> :
R2= 0.461538 (sh 78); R3= 0.626154 (sh 51); R4= 0.634615 (sh 25); R5= 0.583333 (sh 12);
A2= 0.641026 (sh 78); A3= 0.588235 (sh 51); A4= 0.600000 (sh 25); A5= 0.583333 (sh 12);
SYSTEM: Comparison with sample <gatti-11> :
R2= 0.490000 (sh 80); R3= 0.477041 (sh 56); R4= 0.461538 (sh 26); R5= 0.533333 (sh 11);
A2= 0.637500 (sh 80); A3= 0.571429 (sh 56); A4= 0.423077 (sh 26); A5= 0.727273 (sh 11);
SYSTEM: Comparison with sample <gatti-12> :
R2= 0.446341 (sh 81); R3= 0.534024 (sh 52); R4= 0.438095 (sh 29); R5= 0.678571 (sh 15);
A2= 0.604938 (sh 81); A3= 0.423077 (sh 52); A4= 0.586207 (sh 29); A5= 0.400000 (sh 15);
SYSTEM: Comparison with sample <gatti-13> :
R2= 0.455679 (sh 76); R3= 0.564050 (sh 44); R4= 0.445455 (sh 21); R5= 0.500000 (sh 9);
A2= 0.592105 (sh 76); A3= 0.590909 (sh 44); A4= 0.476190 (sh 21); A5= 0.777778 (sh 9);
SYSTEM: Comparison with sample <gatti-14> :
R2= 0.509514 (sh 87); R3= 0.573333 (sh 49); R4= 0.609890 (sh 27); R5= 0.375000 (sh 15);
A2= 0.655172 (sh 87); A3= 0.591837 (sh 49); A4= 0.592593 (sh 27); A5= 0.733333 (sh 15);
SYSTEM: Comparison with sample <gatti-15> :
R2= 0.477733 (sh 77); R3= 0.527410 (sh 46); R4= 0.511111 (sh 19); R5= 0.416667 (sh 7);
A2= 0.688312 (sh 77); A3= 0.630435 (sh 46); A4= 0.368421 (sh 19); A5= 0.428571 (sh 7);
SYSTEM: Mean global distance of new sample from <gatti>: 2.357323
SYSTEM: Second smallest distance is from <gatti> : 2.357323
SYSTEM: Mean distance among <daniele> sample is : 2.585090
SYSTEM: User <daniele> identificated with global mean distance = 2.057170
```

Esempio di una possibile sequenza di stampe di diagnostica

Classificazione

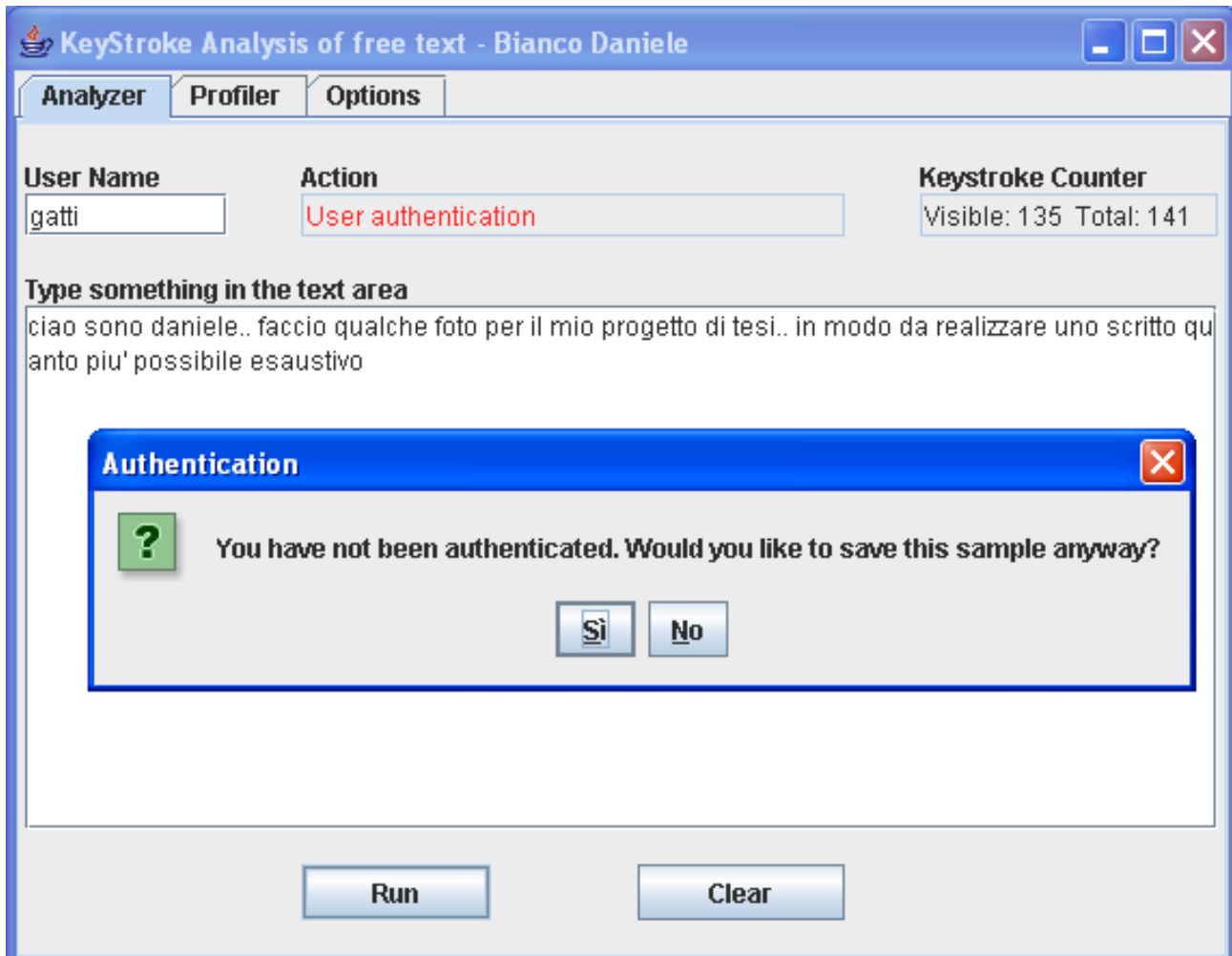
Impostando la costante $k = 1$ nel pannello delle opzioni, il sistema provvederà ad attribuire il campione all'utente che presenta la distanza inferiore al campione provvisto.



Esempio di classificazione

Autenticazione

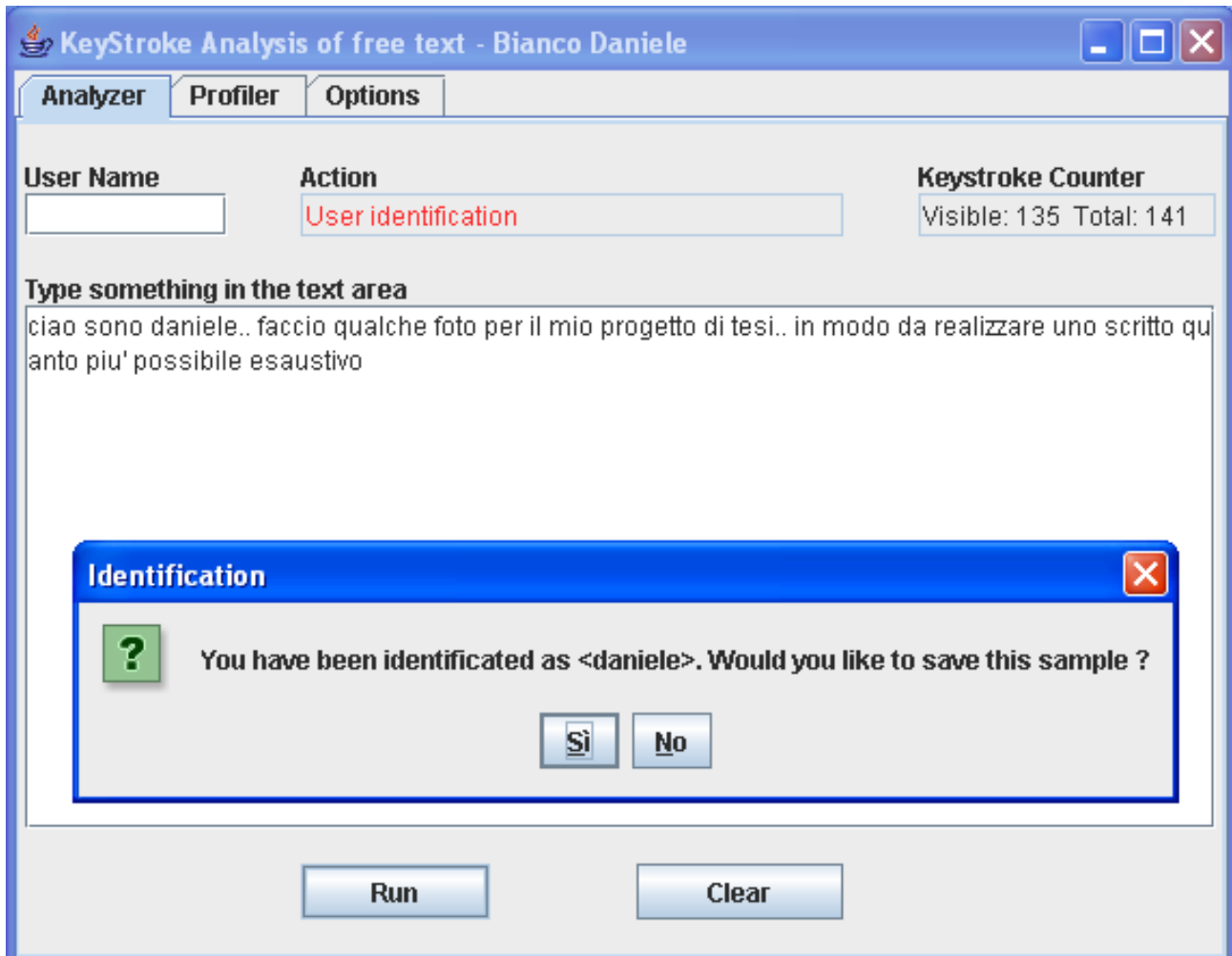
Settando la costante k con un valore diverso da 1, e, inserendo nel campo user-name il proprio nome e nell' area di input un testo libero, il sistema provvedera' all' autenticazione del campione; l' esito puo' essere ovviamente negativo nel caso il campione venga considerato diverso dal profilo utente in questione.



Esempio di Autenticazione

Identificazione

In questa modalita' il sistema provvedera' all' identificazione del campione, ovvero provera' ad attribuire il testo appena inserito ad un utente registrato nel sistema che ha un profilo molto vicino al campione; naturalmente come nel caso dell' autenticazione il campione puo' non venire attribuito a nessuno.



Esempio di Identificazione