



SpaghettiCode

spaghetti.code.g6@gmail.com

Verbale Esterno 2020-12-17

Versione	v1.0.0
Approvazione	Paparazzo Giorgia
Redazione	Masevski Martin
Verifica	Rizzo Stefano
Uso	Esterno
Destinato a	prof. Vardanega Tullio prof. Cardin Riccardo SpaghettiCode Zucchetti S.p.A.

Descrizione

Riassunto dell'incontro realizzato dal gruppo SpaghettiCode tenutosi il 17 Dicembre 2020 in forma di meeting online con Piccoli Gregorio.



Registro delle modifiche

Versione	Nominativo	Ruolo	Data	Descrizione
v1.0.0	Paparazzo Giorgia	Responsabile	2021-01-01	Approvazione del documento
v0.1.0	Rizzo Stefano	Verificatore	2020-12-29	Verifica del documento
v0.0.1	Masevski Martin	Amministratore	2020-12-17	Creazione del documento e prima stesura



Indice

1	Informazioni generali	1
1.1	Informazioni incontro	1
1.2	Riferimenti	1
2	Domande poste	1
3	Resoconto	2
3.1	Necessità di gestire autenticazione e registrazione	2
3.2	Necessità di prevedere formati di file diversi da .csv	2
3.3	Supporto ai browser	2
3.4	Visualizzazione contemporanea di più grafici	2
3.5	Necessità di implementare algoritmi di clustering	2
3.6	Guida introduttiva nell'applicativo	2
3.7	Scopo del progetto	2
3.8	Necessità di salvare il lavoro	2
3.9	Necessità di esportare i grafici	2
3.10	Necessità di effettuare pulizia sui dati forniti	2
3.11	Necessità di dare importanza all'aspetto grafico	2
3.12	Presenza di file o database da cui prendere i dati	3
3.13	Esempio di utilizzo del prodotto finale	3
3.14	Limite alle dimensioni dei dati	3
3.15	Formazione autonoma o verranno forniti materiali di studio	3
3.16	Obbligo sull'utilizzo di JavaScript	3
3.17	Consigli sulle librerie	3
3.18	Presenza di requisiti particolari per il Front-End?	3
3.19	Utilizzo di tipi diversi di grafici per dati diversi	3
4	Conclusione dell'incontro	3



1 Informazioni generali

1.1 Informazioni incontro

- **Luogo:** Applicazione desktop SkypeG;
- **Data:** 2020-12-17;
- **Ora:** 10:15-11:30
- **Partecipanti:**
 - SpaghettiCode
 - Gruppo 5
 - Gruppo 10
 - Gruppo 15
 - Piccoli Gregorio

1.2 Riferimenti

- **link1:** <https://observablehq.com/@mbostock/the-wealth-health-of-nations>;
- **link1:** <https://github.com/mljs/distance>;
- **link1:** <https://cs.stanford.edu/people/karpathy/convnetjs>;
- **link1:** <https://github.com/karpathy/tsnejs>;
- **link1:** <https://github.com/PAIR-code/umap-js>.

2 Domande poste

Vengono riportate le domande che sono state poste nel corso del meeting:

- Necessità di gestire autenticazione e registrazione;
- Necessità di prevedere formati di file diversi da .csv;
- Supporto ai browser;
- Visualizzazione contemporanea di più grafici;
- Necessità di implementare algoritmi di clustering;
- Guida introduttiva nell'applicativo;
- Scopo del progetto;
- Necessità di salvare il lavoro;
- Necessità di esportare i grafici;
- Necessità di effettuare pulizia sui dati forniti;
- Necessità di dare importanza all'aspetto grafico;
- Presenza di file o database da cui prendere i dati;
- Esempio di utilizzo del prodotto finale.



3 Resoconto

3.1 Necessità di gestire autenticazione e registrazione

La gestione della sicurezza non è una priorità: la visualizzazione dei grafici lo è. L'autenticazione viene lasciata a discrezione del gruppo.

3.2 Necessità di prevedere formati di file diversi da .csv

Per recuperare i dati si potrà eseguire query da database o da file, il formato .csv è quello più comune e facile da ottenere. È necessario permettere l'upload diretto di un file.

3.3 Supporto ai browser

La retrocompatibilità con i browser più datati non è un requisito necessario. I browser presi in considerazione sono Firefox e Google Chrome. Viene espresso il desiderio di poter usare il software anche da piattaforme mobile, ma l'implementazione viene lasciata a discrezione del gruppo.

3.4 Visualizzazione contemporanea di più grafici

Il capitolato richiede almeno un grafico per volta. La possibilità di visualizzare insieme più grafici è gradita.

3.5 Necessità di implementare algoritmi di clustering

Viene considerata un'idea interessante.

3.6 Guida introduttiva nell'applicativo

Deve essere presente una guida che aiuti i meno esperti con delle interfacce introduttive o dei percorsi di presentazione.

3.7 Scopo del progetto

Con il progetto si vuole solamente raccogliere delle idee, altrimenti il progetto verrà rimaneggiato.

3.8 Necessità di salvare il lavoro

Viene consigliato di non introdurre questa funzionalità, in quanto renderebbe più complessa la preparazione degli ambienti di test.

3.9 Necessità di esportare i grafici

La funzionalità non è prevista.

3.10 Necessità di effettuare pulizia sui dati forniti

Verranno forniti dati a molte dimensioni anonimizzati. Viene consigliato di cominciare lo sviluppo con dataset facilmente reperibili su Internet. Se ci sarà la necessità di pulire i dataset forniti ci verrà detto più avanti. Verosimilmente i dati saranno da visualizzare as-is. Non si dovrà fare analisi sul contenuto dei dati, ma sulla loro visualizzazione. Ogni possibile scrematura dei dati dovrà essere sempre finalizzata alla visualizzazione.

3.11 Necessità di dare importanza all'aspetto grafico

Viene consigliato di rendere visivamente gradevole il software. Il grafico dovrà necessariamente essere informativo, in secondo luogo potrà essere gradevole alla vista.



3.12 Presenza di file o database da cui prendere i dati

Non sarà necessario collegarsi a database esterni. Saranno forniti file .csv. La presenza di un database interno all'applicativo a scopo interno viene lasciato a discrezione del gruppo.

3.13 Esempio di utilizzo del prodotto finale

Si parte da una query o da un file .csv, la si importa nel sistema il quale mostra i grafici che producono le proiezioni. Quindi a questo punto si ha un grafico in 2D o 3D prodotto a partire da dati con N dimensioni. Se sono contento del grafico bene, altrimenti proseguo con modifiche sulla proiezione dei dati per ottenere una riduzione che sia utile.

3.14 Limite alle dimensioni dei dati

Mediante operazioni di riduzione dimensionale è possibile ridurre il numero di dimensioni. Non sarà necessario effettuare operazioni di riduzione dimensionale. Viene presentato un esempio relativo al grafico scatter plot, che è in grado di includere al massimo 6 dimensioni. le rimanenti (15-6) 9 verranno rappresentate mediante altri grafici.

3.15 Formazione autonoma o verranno forniti materiali di studio

La formazione sarà autonoma, in caso di difficoltà si potrà discutere.

3.16 Obbligo sull'utilizzo di JavaScript

Non ci sono obblighi ma viene sconsigliato l'utilizzo di TypeScript per via della tipizzazione.

3.17 Consigli sulle librerie

ml.js raccoglie molti modi diversi di calcolare le distanze. D3.js permette la visualizzazione ed è caldamente consigliata, tuttavia non è un obbligo. convnet.js fornisce metodi di riduzione dimensionale. Vengono anche mostrati algoritmi per la visualizzazione di cluster come karmathy/tsnejs e umap-js.

3.18 Presenza di requisiti particolari per il Front-End?

Nessuno.

3.19 Utilizzo di tipi diversi di grafici per dati diversi

Bisogna poter visualizzare i dati con vari grafici.

4 Conclusione dell'incontro

Zucchetti anonimizzerà i dati. Il posizionamento dei dati sui vari assi è una scelta che spetta all'utente. Il software potrebbe sottolineare somiglianze tra i dati. È probabile che chi ha fatto il grafico dell'aspettativa di vita, preso d'esempio durante le domande, abbia fatto delle prove per evidenziare certe proprietà. Ad esempio, ci siamo resi conto che la popolazione non influenza molto il PIL.

Zucchetti chiuderà per Natale, quindi non saranno reperibili fino al loro rientro.