

UNIVERSITÀ DEGLI STUDI DI PADOVA



SCUOLA DI SCIENZE

CORSO DI LAUREA IN INFORMATICA

---

## Piano di Lavoro

---

*Studente:*

**Eleonora SIGNOR - 1049099**

*Azienda:*

**Zucchetti S.p.A**

10 maggio 2019

## Contatti

**Studente:** Eleonora Signor, [eleonora.signor@studenti.unipd.it](mailto:eleonora.signor@studenti.unipd.it), + 39 346 41 37 765

**Tutor aziendale:** Gregorio Piccoli, [gregorio.piccoli@zucchetti.it](mailto:gregorio.piccoli@zucchetti.it), + 39 0371 59 457 11

**Azienda:** Zucchetti S.p.A, Via Giovanni Cittadella, 735137 - Padova, <https://www.zucchetti.it>

## Scopo dello stage

Zucchetti S.p.A è la prima software house italiana per storia e dimensione. Studia e realizza soluzioni software gestionali, hardware e servizi innovativi per soddisfare le specifiche esigenze di:

- Aziende di qualsiasi settore e dimensione, banche e assicurazioni;
- Professionisti (commercialisti, consulenti del lavoro, avvocati, curatori fallimentari, notai ecc.), associazioni di categoria e CAF;
- Pubblica amministrazione locale e centrale (Comuni, Province, Regioni, Ministeri, società pubbliche ecc.).

Lo stage consiste in un periodo di formazione dello studente "sul campo" assistito in azienda, mirato all'apprendimento per mezzo di compiti reali e sperimentazione diretta dei processi e contesti lavorativi. Il progetto di stage da sviluppare consiste nello sviluppo di un *Reticolo della Conoscenza* in grado di mappare e correlare tra loro tutte le possibili domande che potranno essere poste in fase di colloquio aziendale ad un candidato, in modo da agevolarne ed ottimizzarne i risultati. Lo studente, per la sua realizzazione, dovrà seguire i seguenti passi:

- Sviluppo di una Rete neurale con configurazione ad "autoencoder" e apprendimento della stessa mediante training set ottenuto dalle risposte date finora dai candidati ai colloqui di lavoro;
- Valutazione probabilistica delle risposte ottenute dai candidati durante i colloqui con l'azienda;
- Analisi dei dati mediante tecniche lineari (Principal Component Analysis) e confronto con quanto riscontrato dalla Rete neurale;
- Rappresentazione dei risultati finali ottenuti su un Reticolo della Conoscenza.

## Interazione tra studente e tutor aziendale

Durante l'intero periodo di stage lo stagista lavorerà a stretto contatto con il tutor aziendale Gregorio Piccoli e con il personale aziendale, in modo da favorire l'integrazione dello studente nell'ambiente lavorativo e garantire il raggiungimento degli obiettivi prefissati entro i termini pattuiti. Per verificare lo stato di avanzamento del lavoro svolto verrà effettuato una revisione dello stesso in modo da chiarirne eventualmente gli obiettivi, affinare la ricerca e aggiornare il Piano di Lavoro, se necessario.

## **Contenuti formativi previsti**

Durante questo progetto di stage lo studente avrà occasione di approfondire le sue conoscenze:

- Nelle aree dell'Intelligenza artificiale mediante:
  - Lo studio e lo sviluppo di una Rete neurale;
  - L'utilizzo di tecniche di Apprendimento automatico per un sistema supervisionato.
- Nell'Analisi dei dati con l'utilizzo:
  - Del linguaggio R;
  - Di tecniche di semplificazione del dato.

## Pianificazione del lavoro

### Pianificazione settimanale

- **Prima Settimana - Studio Reti neurali e tecniche di Apprendimento automatico**

- Studio delle Reti neurali;
- Studio della tecnica di Autoencoder.

Durante questa settimana lo studente verrà prima introdotto alla struttura amministrativa da Zucchetti e successivamente rivolgerà parte del monte delle ore previste per l'apprendimento dei supporti hardware e software utili per la realizzazione del progetto, inerenti la parte di AI. Lo studente sarà chiamato ad effettuare uno studio approfondito del significato, delle implicazioni e delle modalità di realizzazione di una Rete neurale con studio della tecnica di Autoencoder per la minimizzazione dei dati.

- **Seconda Settimana - Analisi e definizione soluzione software del problema con tecniche di AI**

- Realizzazione Rete neurale;
- Addestramento dati dei test.

Durante questa settimana lo studente avrà il compito di procedere con la realizzazione di una Rete neurale mediante l'applicazione della tecnica di Autoencoder, con l'obiettivo prima di generare nuovi dati comprimendo l'input su uno spazio di variabili latenti e successivamente ricostruirne l'output sulla base delle informazioni acquisite. Inoltre lo studente avrà il compito di prelevare i dati generati dai colloqui ed utilizzarli per addestrare la Rete neurale. Il training set consisterà in un vettore di input composto dalle domande ai quesiti a cui verrà associata una determinata classificazione in base alle risposte ottenute.

Gli "autoencoder" potranno assumere un numero molteplici di forme e nel caso di risultati che evidenzino situazioni contrastanti si procederà verso un'analisi dei casi e allo studio e sperimentazione di ulteriori tecniche di addestramento della Rete. Per procedere alla realizzazione della Rete neurale si farà uso di uno tra le seguenti librerie: ConvnetJS in linguaggio JavaScript o DeepLearning4J in linguaggio Java, sulle quali verrà effettuata preventivamente un'analisi per valutarne la convenienza d'uso.

- **Terza Settimana - Valutazione test**

- Valutazione probabilistica delle risposte ottenute durante l'esecuzione dei test.

Durante questa settimana lo studente avrà il compito di valutare per ogni quesito le risposte ottenute da ogni candidato ed effettuarne una stima.

- **Quarta Settimana - Test e documentazione**

- Durante questa settimana, lo stagista dovrà documentare e testare il codice prodotto durante le settimane precedenti.

- **Quinta Settimana - Studio tecniche lineari**

- Studio del linguaggio R;
- Studio della tecnica Principal Component Analysis (PCA).

Durante questa settimana lo studente dovrà effettuare lo studio del linguaggio di programmazione R e del suo ambiente di sviluppo specifico in funzione dell'analisi dei dati. Inoltre avrà il compito di apprendere e approfondire l'uso della tecnica Principal Component Analysis (PCA) da impiegare nel campo della statistica multivariata come approccio per la semplificazione dei dati.

• **Sesta Settimana - Analisi e definizione soluzione software del problema con tecniche lineari**

- Applicazione Principal Component Analysis (PCA) ai risultati dei test;
- Confronto risultati ottenuti dalle tecniche Autoencoder e Principal Component Analysis (PCA).

Durante questa settimana la studente dovrà effettuare l'applicazione della tecnica di PCA che avverrà mediante trasformazione lineare dei risultati dei test. Si procederà dapprima con l'individuazione delle variabili coinvolte dalla trasformazione lineare e con la proiezione delle stesse su un nuovo sistema cartesiano per mezzo della trasformazione stessa. Successivamente verranno individuate quali sono le variabili con maggiore varianza e verranno proiettate in modo decrescente su ogni asse secondo un ordine crescente. La riduzione della complessità avverrà mediante analisi delle prime variabili coinvolte.

Una volta effettuata la risoluzione i risultati ottenuti dovranno venire confrontati con quanto rilevato durante la terza settimana con lo scopo di effettuarne delle considerazioni e stime.

• **Settima Settimana - Realizzazione Reticolo della Conoscenza**

- Studio modalità e rappresentazione dei risultati in Reticolo della Conoscenza.

Durante questa settimana lo studente dovrà assolvere alla mansione di studio del significato e delle modalità di sviluppo su casi concreti di un Reticolo della Conoscenza. Successivamente dovrà rivolgere le proprie ore di lavoro verso lo sviluppo del caso concreto, oggetto dello stage, con l'impiego dei risultati ottenuti nelle settimane precedenti. L'obiettivo che verrà perseguito dallo studente consisterà in una mappatura completa e performante dei quesiti da svolgere durante il test di un candidato all'assunzione.

• **Ottava Settimana - Test e Documentazione**

- Durante questa settimana, lo stagista dovrà documentare e testare il codice prodotto durante le settimane precedenti.

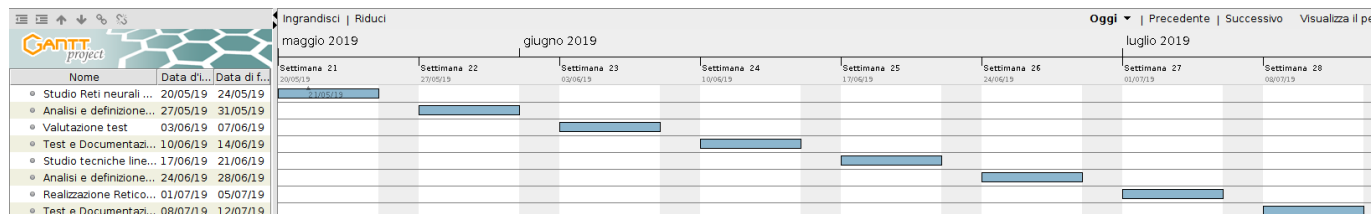
## Ripartizione ore

La pianificazione, in termini di quantità di ore di lavoro, sarà così distribuita:

| Durata in ore     | Date (inizio - fine)    | Descrizione dell'attività  |
|-------------------|-------------------------|--|
| 40                | 20/05/2019 - 24/05/2019 | Studio Reti neurali e tecniche di Apprendimento automatico                 |
| 40                | 27/05/2019 - 31/05/2019 | Analisi e definizione soluzione software del problema con tecniche di AI   |
| 40                | 03/06/2019 - 07/06/2019 | Valutazione test   |
| 40                | 10/06/2019 - 14/06/2019 | Test e Documentazione  |
| 40                | 17/06/2019 - 21/06/2019 | Studio tecniche lineari  |
| 40                | 24/06/2019 - 28/06/2019 | Analisi e definizione soluzione software del problema con tecniche lineari |
| 40                | 01/07/2019 - 05/07/2019 | Realizzazione Reticolo della Conoscenza                                    |
| 40                | 08/07/2019 - 12/07/2019 | Test e Documentazione  |
| <b>Totale ore</b> | <b>320</b>              |  |

## Diagramma di Gantt

Di seguito è riportato il diagramma di Gantt relativo al piano di lavoro previsto.



## Obiettivi

### Notazione

Si farà riferimento ai requisiti secondo le seguenti notazioni:

- **O** per i requisiti obbligatori, vincolanti in quanto obiettivo primario richiesto dal committente;
- **D** per i requisiti desiderabili, non vincolanti o strettamente necessari, ma dal riconoscibile valore aggiunto;
- **F** per i requisiti facoltativi, rappresentanti valore aggiunto non strettamente competitivo.

Le sigle precedentemente indicate saranno seguite da una coppia sequenziale di numeri, identificativo del requisito.

### Obiettivi fissati

Si prevede lo svolgimento dei seguenti obiettivi:

- **Obbligatori**
  - *O01*: Studio e analisi dei sistemi coinvolti;
  - *O02*: Realizzazione della Rete neurale;
  - *O03*: Addestramento della Rete neurale;
  - *O04*: Analisi dei dati dei test e stima probabilistica delle risposte mediante l'uso della Rete neurale;
  - *O05*: Realizzazione Reticolo della Conoscenza;
  - *O06*: Produzione di una documentazione completa come resoconto delle attività svolte e delle osservazioni effettuate nel contesto della Rete neurale;
  - *O07*: Rispetto delle regole aziendali vigenti e delle norme in uso in materia di igiene, sicurezza e salute presso i luoghi di lavoro;
  - *O08*: Mantenimento della riservatezza necessaria nei confronti dei dati osservati durante lo svolgimento dello stage e per quanto riguarda i processi di produzione e/o dei prodotti dell'azienda ospitante.
- **Desiderabili**
  - *D01*: Analisi dei dati dei test mediante tecnica di PCA;
  - *D02*: Valutazione e confronto dei risultati ottenuti dalla Rete neurale e dalla PCA.
  - *D03*: Sperimentazione di un numero molteplici di architetture sulla Rete neurale;
  - *D04*: Realizzazione della necessaria documentazione che attesti i risultati ottenuti con la tecnica di PCA e ne valuti i pregi e i difetti nei confronti dell'uso della Rete neurale.
- **Facoltativi**
  - *F01*: Realizzazione automa in grado di imparare dalla percentuale di risposte esatte ai test;
  - *F02*: Realizzazione della necessaria documentazione che attesti la realizzazione e le metodologie di apprendimento dell'automa.

## Prodotti attesi a fine stage

### Obbligatori

1. Realizzazione di una Rete neurale;
2. Addestramento della Rete neurale sulla base dei dati dei test;
3. Produzione documentazione relativa alla Rete e ai test;
4. Realizzazione di un Reticolo della Conoscenza sulla base delle analisi effettuate nei test.

### Desiderabili

1. Analisi dei test sulla base della metodologia PCA;
2. Documentazione relativa alla tecnica PCA.

### Opzionali

1. Realizzazione di un automa in grado di autoapprendere le risposte dei test;
2. Produzione della documentazione necessaria che permetta la mantenibilità dell'automa.



---

## Approvazione

Il presente piano di lavoro è stato approvato dai seguenti

---

Gregorio Piccoli

Tutor aziendale

---

Eleonora Signor

Stagista

---

Prof. Tullio Vardanega

Tutor interno

Data