# Traccia 1

**Traccia Esercitazione: Generatore di Testo per Creare Poesie o Frasi Creative**

**Obiettivo:**

Gli studenti dovranno usare il modello RNN (sia la versione manuale che quella implementata con Keras) per creare un generatore di testo che possa essere applicato a un caso d'uso creativo. L'idea principale è di costruire un modello che generi poesie, frasi o brevi racconti a partire da un testo di "seed" dato.

**Fase 1: Preparazione dei Dati**

Gli studenti devono raccogliere un corpus di testo da utilizzare per addestrare il modello. Ecco alcune opzioni che potrebbero esplorare:

* **Poesie di autori famosi** (ad esempio, poesie di Emily Dickinson, Shakespeare, etc.)
* **Testi di libri o racconti** (ad esempio, opere di scrittori come J.K. Rowling, J.R.R. Tolkien, etc.)
* **Dialoghi di film o serie TV** (può essere interessante per generare conversazioni creative)
* **Descrizioni di paesaggi, situazioni o eventi** (per generare descrizioni evocative o "scene" narrative)

In questo modo, gli studenti dovranno:

1. **Raccogliere un corpus di testo**: Possono utilizzare siti web come Project Gutenberg per scaricare opere di dominio pubblico.
2. **Pulire e Preprocessare il testo**: Simile a quanto fatto nel codice che hai fornito, rimuovendo caratteri speciali, tokenizzando il testo, e creando un vocabolario.

**Fase 2: Modello di Addestramento**

Dopo aver preparato il corpus, gli studenti devono:

1. **Addestrare un modello RNN** come nel codice che hai scritto.
2. **Esplorare l'ottimizzazione del modello**: Possono provare a modificare vari parametri come:
   * Numero di unità nel layer RNN.
   * Lunghezza delle sequenze in input (modificare seq\_length).
   * Numero di epoche di addestramento.
   * Tipologie di ottimizzatori o funzioni di attivazione.
3. **Valutazione**: Gli studenti devono osservare le metriche di addestramento, come la **loss** e l'**accuracy**, e analizzare se il modello sta effettivamente imparando a generare testo coerente.

**Fase 3: Generazione di Testo Creativo**

Una volta addestrato il modello, gli studenti devono:

1. **Generare testo a partire da un "seed"**: Devono inserire una frase di partenza (ad esempio: "Un albero inizia a crescere", "Nel buio della notte", etc.) e il modello genererà una continuazione.
2. **Sperimentare con diverse temperature**: La temperatura regola il livello di creatività e casualità nel testo generato. Gli studenti dovrebbero testare valori di temperatura diversi (ad esempio 0.5, 1.0, 1.5) e analizzare come cambia il testo generato.

**Fase 4: Analisi e Miglioramenti**

1. **Analisi del testo generato**: Gli studenti devono valutare la qualità del testo generato. È sensato? È coerente con il contesto dato nel "seed"? Riusciresti a riconoscere un tema comune o un pattern?
2. **Miglioramenti e ottimizzazioni**:
   * **Aggiunta di un layer LSTM o GRU**: Per migliorare le performance della RNN, gli studenti possono provare a sostituire la SimpleRNN con una LSTM o una GRU, che sono generalmente più efficaci nel catturare le dipendenze a lungo termine.
   * **Augmentare il dataset**: Utilizzare tecniche di data augmentation sul testo, come la generazione di frasi simili o l'inclusione di parole con sinonimi, per arricchire il modello.

**Fase 5: Presentazione del Progetto**

Ogni studente o gruppo di studenti dovrà presentare:

1. **Il modello e l'addestramento**: Spiegare come è stato addestrato il modello, con i parametri utilizzati e la strategia adottata.
2. **Risultati della generazione**: Mostrare alcuni esempi di testo generato dal modello, sia con temperatura bassa che alta.
3. **Criticità e riflessioni**: Condividere i punti di forza e di debolezza del modello, le difficoltà incontrate e le possibili migliorie.

# Traccia 2

**Fase 1: Preparazione dei Dati**

Gli studenti devono seguire lo stesso processo della Traccia 1, ma con un focus sull'analisi dei risultati che una LSTM potrebbe fornire rispetto a una RNN semplice.

1. **Raccogliere un corpus di testo** da utilizzare per addestrare il modello. Gli studenti possono scegliere:
   * Poesie di autori famosi (come nella traccia 1).
   * Testi di racconti o romanzi (come nei libri di J.K. Rowling o J.R.R. Tolkien).
   * Dialoghi di film o serie TV.
   * Descrizioni di paesaggi o situazioni evocative.
2. **Pulire e preprocessare il testo**, rimuovendo caratteri speciali, tokenizzando e creando un vocabolario. Qui, gli studenti dovrebbero concentrarsi anche sull'analisi delle sequenze di lunghezza maggiore, perché le LSTM gestiscono meglio le dipendenze a lungo termine, quindi il dataset potrebbe essere più ampio rispetto a una RNN.

**Fase 2: Costruzione e Addestramento del Modello LSTM**

1. **Creare il modello LSTM**:
   * Utilizzare una **LSTM** invece di una **RNN**. Gli studenti devono comprendere le differenze tra le due architetture e come una LSTM possa migliorare la generazione di testo rispetto a una RNN semplice.
   * Modificare il numero di unità nel layer LSTM.
   * Impostare la **lunghezza delle sequenze** (in modo simile alla traccia precedente, ma le LSTM potrebbero trarre vantaggio da sequenze più lunghe).
2. **Sperimentare con l'ottimizzazione**:
   * Modificare **parametri come il numero di epoche**, **learning rate**, **ottimizzatori**, e **funzioni di attivazione**.
   * **Valutare la perdita** e l'accuratezza del modello, ma concentrarsi maggiormente sulla qualità del testo generato.
3. **Comportamento dell'algoritmo**:
   * **LSTM vs RNN**: Gli studenti devono osservare come la LSTM si comporta rispetto alla RNN nel catturare relazioni a lungo termine e nel generare un testo più coerente e fluido.

**Fase 3: Generazione di Testo Creativo**

1. **Generare il testo** a partire da un "seed" dato, come nelle tracce precedenti:
   * La frase di partenza potrebbe essere un testo evocativo come: "Una stella cadente", "Nel cuore del bosco", etc.
2. **Sperimentare con la temperatura**:
   * Come nella traccia 1, gli studenti dovrebbero testare diverse **temperature** (0.5, 1.0, 1.5) per controllare il livello di creatività nel testo generato. La temperatura bassa rende il testo più prevedibile, mentre quella alta lo rende più casuale e fantasioso.

**Fase 4: Analisi e Miglioramenti**

1. **Analisi del testo generato**:
   * Gli studenti dovranno esaminare la qualità del testo generato. Alcune domande da considerare:
     + **Coerenza**: Il testo generato è coerente? Ha una narrazione sensata?
     + **Creatività**: Il modello riesce a produrre frasi poetiche o evocative?
     + **Fluidità**: La LSTM ha migliorato la capacità del modello di generare sequenze più lunghe e coerenti rispetto a una RNN?