

Mini-Challenge 1: Hyperparameter und Model Tuning

Ziel

Implementiere, trainiere und validiere ein Neuronales Netzwerk Modell für einen Klassifikationstask mit geeignetem Dataset. Der zu modellierende Task und die dazu passenden Daten sollen am Anfang der Mini-Challenge individuell vereinbart werden. Dabei können die Studierenden auch ihre Wünsche einbringen und Task/Datensatz vorschlagen. Dabei ist aber darauf zu achten, dass der Fokus auf das Training und die Evaluation gelegt und nicht zu viel Zeit mit Datenbeschaffung, -Analyse und -Vorverarbeitung verbraucht wird. Der Datensatz sollte also nicht zu gross sein, um viele verschiedene Modell- und Parametervarianten evaluieren zu können. Ausserdem sollten nicht zu ausgeklügelte komplexe Modell-Architekturen untersucht werden. Ein MLP für die Klassifikation allenfalls mit ein paar CNN-Layern für Feature-Extraktion sollte ausreichen.

Arbeitsschritte

Schritt 1: Auswahl Task / Datensatz

1. Mache Dir Gedanken, mit welchen Daten Du arbeiten möchtest und welcher Task gelernt werden soll.
2. Diskutiere die Idee mit dem Fachcoach.

--> **Absprache/Beschluss** mit Fachcoach über Daten/Task.

Schritt 2: Daten Kennenlernen

1. Mache Dich mit dem Datensatz vertraut, indem Du eine explorative Analyse der Features durchführst: z.B. Vergleich der Klassen pro Feature, Balanciertheit der Klassen.
2. Führe ein geeignetes Preprocessing durch, z.B. Normalisierung der Daten.

Schritt 3: Aufbau Modellierung

1. Lege fest, wie (mit welchen Metriken) Du die Modelle evaluieren möchtest. Berücksichtige auch den Fehler in der Schätzung dieser Metriken.
2. Implementiere Basisfunktionalität, um Modelle zu trainieren und gegeneinander zu evaluieren. Wie sollen die Gewichte initialisiert werden?

--> **Absprache/Beschluss** mit Fachcoach über die zu untersuchenden Varianten (Schritte 4)

Schritt 4: Evaluation

Bei der Evaluation ist darauf zu achten, dass das Vorgehen stets möglichst reflektiert erfolgt und versucht wird, die Ergebnisse zu interpretieren. Am Schluss soll auch ein Fazit gezogen werden, darüber welche Variante am besten funktioniert.

1. Training mit SGD, ohne REG, ohne BN:

Untersuche verschiedene Modelle unterschiedlicher Komplexität, welche geeignet sein könnten, um das Klassifikationsproblem zu lösen. Verwende Stochastic Gradient Descent - ohne Beschleunigung, ohne Regularisierung (REG) und ohne Batchnorm (BN).

- a. Für jedes Modell mit gegebener Anzahl Layer und Units pro Layer führe ein sorgfältiges Hyper-Parameter-Tuning durch (Lernrate, Batch-Grösse). Achte stets darauf, dass das Training stabil läuft. Merke Dir bei jedem Training, den Loss, die Performance Metrik(en) inkl. Schätzfehler, die verwendete Anzahl Epochen, Lernrate und Batch-Grösse.
- b. Variiere die Anzahl Layer und Anzahl Units pro Layer, um eine möglichst gute Performance zu erreichen. Falls auch CNNs (ohne Transfer-Learning) verwendet werden variiere auch Anzahl Filter, Kernel-Grösse, Stride, Padding.
- c. Fasse die Ergebnisse zusammen in einem geeigneten Plot, bilde eine Synthese und folgere, welche Modell-Komplexität Dir am sinnvollsten erscheint.

2. Nutzen der Regularisierung

Ziehe nun verschiedene Regularisierungsmethoden bei den MLP Layern in Betracht:

- a. L1/L2 Weight Penalty
- b. Dropout

Evaluieren Sie den Nutzen der Regularisierung, auch unter Berücksichtigung verschiedener Regularisierungsstärken.

Beschreiben Sie auch kurz, was allgemein das Ziel von Regularisierungsmethoden ist (Regularisierung im Allgemeinen, sowie auch Idee der einzelnen Methoden).

Inwiefern wird dieses Ziel im gegebenen Fall erreicht?

3. Nutzen von Batchnorm BN (ohne REG, mit SGD)

Evaluieren Sie, ob Batchnorm etwas bringt. Beschreiben Sie kurz, was die Idee von BN ist, wozu es helfen soll.

4. Nutzen von Adam (ohne BN, ohne / mit REG)

Evaluieren Sie, ob Sie mit Adam bessere Resultate erzielen können.

Schritt 5: Präsentation, Bericht

1. **Präsentation** (~10m): Kurze Präsentation mit Diskussion der wichtigsten Ergebnisse.
Q&A (~10min): Klärung von Verständnisfragen zu Stochastic Gradient Descent, Parameter Tuning, Regularisierung, Batchnorm und Optimizers.
2. **Bericht** in Form eines gut dokumentierten, übersichtlichen Jupyter Notebooks. Dieses soll schliesslich auch abgegeben werden und dem Fachexperten erlauben, die Schritte nachzuvollziehen (allenfalls auch das Training erneut laufen zu lassen).

Zeitlicher Rahmen:

Wird beim Schritt 1 verbindlich festgelegt.

Beurteilung

Beurteilt wird auf Basis des abgegebenen Notebooks:

- Vollständige und korrekte Umsetzung der vereinbarten Aufgabestellung.
- Klare, gut-strukturierte Umsetzung.
- Schlüssige Beschreibung und Interpretation der Ergebnisse. Gut gewählte und gut kommentierten Plots und Tabellen.

- Vernünftiger Umgang mit (Computing-)Ressourcen.
- Verständliche Präsentation der Ergebnisse.