

Mini-Challenge VDL: Generative Modelle und Attention

Ziel

Vertieftes Verständnis und praktische Umsetzung von generativen Modellen eingebautem Attention-Mechanismus, möglicherweise in Form einer Transformer-Architektur. Die konkreten Arbeiten für diese MC sollen mit dem Fachcoach vorbesprochen werden.

Typische Varianten:

- Text-Generierung: implementiere, trainiere und validiere ein Deep Learning Modell für Image Captioning. Beim Generieren der Sequenz von Wörtern für die Caption soll mit Hilfe von Attention gelernt werden, auf verschiedene Bereiche im Bild zu achten.
- Bild-Generierung:

 implementiere, trainiere und validiere ein Deep Learning Modell für die Generierung von Bildern, bedingt auf Eigenschaften, welche die auf dem Bild ersichtlichen Inhalte erfüllen sollen (z.B. Objekte von einem gewissen Typ, in einer gewissen Grösse, an einer bestimmten Stelle mit gegebener Orientierung im Bild zu finden, allenfalls basierend auf einer textuellen Beschreibung. Hier könnte der Attention-Mechanismus auf Teile der Bedingung achten.
 Alternativ kann in dieser Variante die Wirkung von Attention separat anhand eines Toy-Beispiels untersucht werden.

Arbeitsschritte

Schritt 1: Konkretisierung des Tasks

Hier wird festgelegt, welcher Task konkret gelöst werden soll. Zwei typische Beispiel-Settings sind oben beschrieben. Hier soll aber weiter ins Detail gegangen werden und konkretisiert werden, welcher Task gelöst werden soll, welche Daten genommen werden und welche Architekturen in Betracht gezogen werden sollen.

Treffen mit Fachexperten: Diskussion und Beschluss zum Vorgehen.

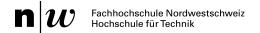
Schritt 2: Vorbereiten der Daten und Aufbau Modellierung

Verschaffe Dir einen Überblick über die Daten (wenig explorative Datenanalyse) und erstelle ein geeignetes Datenset (pre-processing falls noch notwendig), damit die Modelle damit trainiert werden können.

Überlege Dir konkret, welche Modell-Architekturen Du umsetzen möchtest. Hier sollen mindestens zwei Varianten aufgebaut werden – eine Baseline (z.B. ohne Attention) und eine Variante. Überlege auch, welche Metriken für die Evaluation herangezogen werden sollen. Baue nun die Modelle und alle fürs Training und die Evaluation notwendige Funktionalität (inkl. Loss, Metriken, Anbindung an W&B, etc.).

Führe dann für beide Modell-Varianten je einen Overfitting-Test durch (typischerweise mit einem Mini-Batch). Stell sicher, dass in jedem Update-Schritt dieselben Daten verwendet werden.

Mache Dir nun Gedanken, mit welchen Einstellungen (Hyperparameter, Modell-Spezifikationen) Du trainieren und evaluieren möchtest.



Treffen mit Fachexperten: Aufzeigen der Overfitting-Ergebnisse; Beschluss, mit welchen Einstellungen trainiert und evaluiert werden sollen.

Schritt 3: Training und Evaluation

Trainiere und evaluiere die vereinbarten Modelle mit Einstellungen. Analysiere die Ergebnisse, bringe sie in Verbindung mit den Modellen und den verwendeten Settings.

Schritt 4: Präsentation / Bericht

Präsentation (~10m): Präsentation mit Diskussion der wichtigsten Ergebnisse.

Q&A (~10min): Klärung von Verständnisfragen zu Theorie und Umsetzung.

Bericht in Form eines gut dokumentierten, übersichtlichen Jupyter Notebooks.

Dieses soll schliesslich auch abgegeben werden und dem Fachexperten erlauben, die Schritte nachzuvollziehen (ohne, dass das Training erneut laufen gelassen werden muss).

Zeitlicher Rahmen:

Wird beim Schritt 1 verbindlich festgelegt.

Beurteilung

Beurteilt wird auf Basis des abgegebenen Notebooks und der Präsentation inkl. Q&A:

- 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.