Bearbeitung und Besprechung von Arbeitsblatt 3

Die Schülerinnen und Schüler erstellen zum ersten Mal einen mehrstufigen Entscheidungsbaum ohne detaillierte Anweisungen der Lehrkraft, die nur technische Anweisungen zur Bedienung des Tools gibt (Video\_Basisfunktionen\_EB). Insbesondere gibt es keine Anweisung, wie man Prädiktorvariablen auswählt oder wann der Erstellungsprozess gestoppt werden soll. Dabei können Schülerinnen und Schüler verschiedene Ansätze wählen.

Einige Schülerinnen und Schüler probieren nur zufällige Variablen aus, andere argumentieren inhaltlich **kontextbasiert** bei der Auswahl der Variablen und wieder andere erstellen und verwenden **datenbasiert** formale Kriterien. Manche beenden den Erstellungsprozess relativ schnell, wenn sie keine Verbesserung mehr sehen, andere versuchen den Baum so lange zu erweitern bis kein Fall im Trainingsdatensatz mehr falsch klassifiziert wird (Abbruchkriterium: **Null Fehler**). All diese Ansätze sind zu diesem Zeitpunkt und für diese Aufgabenstellung legitim. Die spätere Besprechung soll mögliche Strategien aufgreifen und zu einer Diskussion führen, die weitere Lernanlässe initiiert.

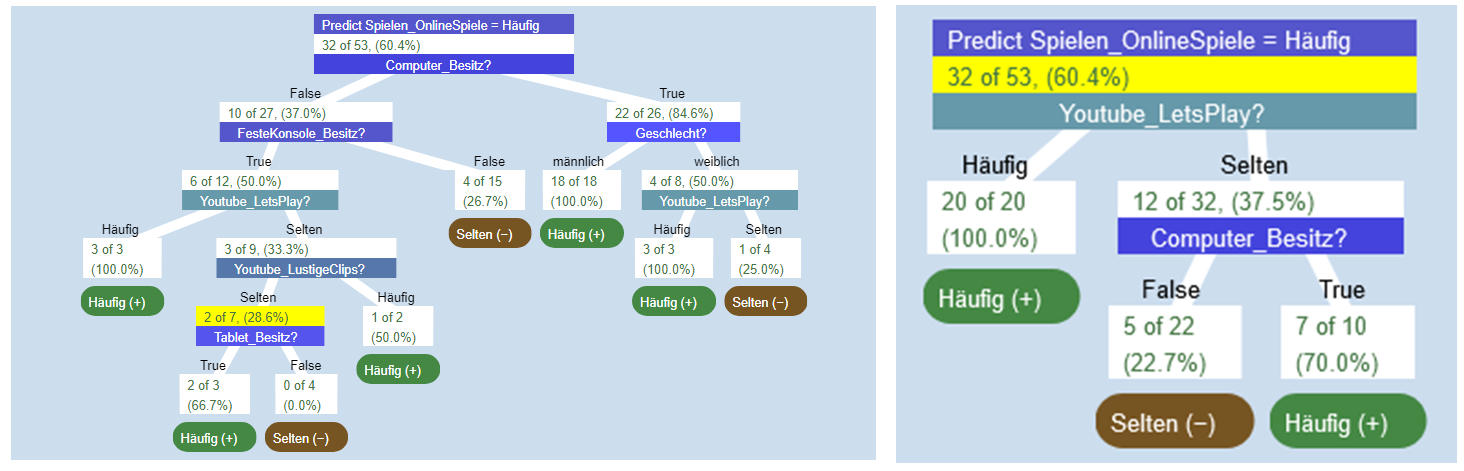
Es können anschließend folgende Fragen aufgeworfen werden:

1. Wie werden solche Entscheidungsbäume bei Onlineplattformen erstellt? (**kontextbasiert** vs. **datenbasiert**)
2. Welche Entscheidungsbäume funktionieren für neue Personen besser, kleine Bäume oder große Bäume? (Abbruchkriterium: **Null Fehler** vs. andere)

Aus den Fragen können folgende Lernanlässe hergeleitet werden:

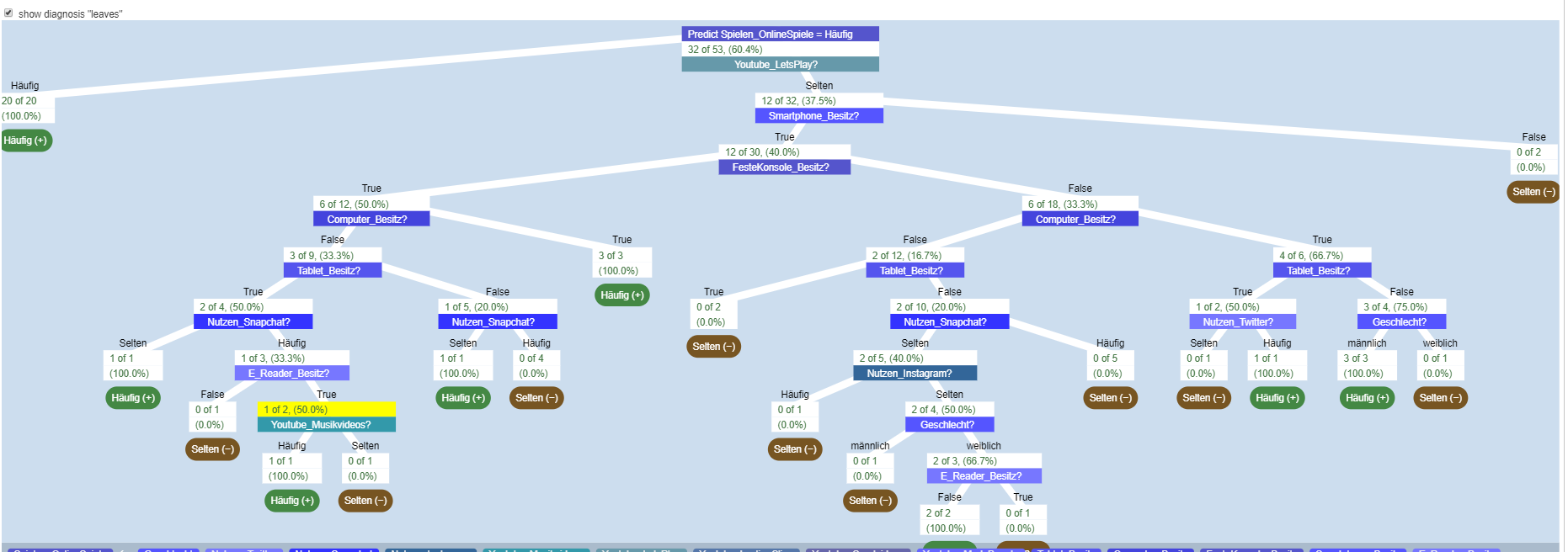
1. Wir können uns anschauen, wie Bäume algorithmisch erstellt werden, damit sie durch einen Computer erstellt werden können. Anschließend können wir die Vorgehensweise von Computern mit unseren eigenen „menschlichen“ Vorgehensweisen gegenüberstellen.
2. Wir können herausfinden, ob große oder kleine Bäume besser funktionieren, indem wir Testdaten verwenden

**Beispiele: Schülerergebnisse**



Der linke Baum enthält weitaus mehr Variablen als der rechte Baum, schneidet jedoch nur geringfügig besser ab als der rechte Baum (7 Fehlklassifikationen gegenüber 8 Fehlklassifikationen). Während der Präsentation argumentierte der Student, der den linken Baum erstellt hat, für die Plausibilität der verschiedenen Variablen und erklärte kontextbezogen, warum die verschiedenen Variablen geeignet sein könnten, die Zielvariable vorherzusagen. Der Schüler wählte zunächst die plausiblen Variablen durch intuitives Argumentieren aus und probierte dann verschiedene Kombinationen aus.

Der Schüler, der den rechten Baum erstellte, argumentierte auf eine andere Weise. Er testete verschiedene Variablen über "Trial and Error" als oberste Entscheidungsregel und suchte nach Teildatensätzen, bei denen die relative Häufigkeit des Zielwertes "nahe 100 % oder nahe 0 %" ist, um zu ermitteln, welche Variablen als Prädiktoren für die Zielvariable geeignet sein könnten. Außerdem gab er an, dass er den Wachstumsprozess sehr früh stoppte, damit die finalen Teildatensätze "repräsentativ" bleiben. Zusammenfassend lässt sich sagen, dass der Schüler des linken Baumes intuitiv entschied, wie es Menschen normalerweise tun, während der Schüler des rechten Baumes sich eher "wie eine Maschine" ohne Kontextwissen verhielt und den später zu thematiosierenden Entscheidungsbaum-Algorithmus antizipierte.



Der Schüler, der diesen Baum erstellt hat, hat die ersten Knoten nach inhaltlicher Plausibilitiät und Fehlklassifikationsrate ausgewählt. In den tieferen Stufen war es nur noch ein wildes ausprobieren aller möglichen Variablen. Allerdings hatte der Schüler das Ziel am Ende null Fehlklassifikationen zu haben.

Wir halten es für wünschenswert, dass die Schüler mit vertretbarem Aufwand manuell einen Baum erstellen können, der die Trainingsdaten "perfekt" klassifiziert. Die Erfahrungen aus dieser Phase wollen wir später nutzen, um eine Diskussion über Overfitting anzuregen. Das Testen eines solchen "perfekten" Baums und die Feststellung, dass er bei Testdaten möglicherweise schlechter abschneidet, ist wichtig, um eine Diskussion über Overfitting der Trainingsdaten und Pruning zu motivieren. Die Erfahrung, dass man in den unteren Stufen nur wild ausprobiert hat, um die letzten Fehklassifikationen noch auszumerzen macht es Plausibel, dass die so erstellten regeln nicht für eine Vorhersage taugen und somit zu Overfitting führen.