

Effiziente Klassifikation und Ranking mit paarweisen Vergleichen

Zwischenbericht

Sang-Hyeun Park

Betreuer: Prof. Fürnkranz
Technische Universität Darmstadt
Fachbereich Informatik
Knowledge Engineering

13. Juli 2006



Gliederung

Einleitung

(Multi)klassifikation

Binärisierung

Round Robin

Effiziente Klassifizierung

Dekodierung

Quick Weighted Voting

Effizientes Ranking

Ranking

Dekodierung

„Schweizer-System“

Ausblick

weitere Ansätze



Gliederung

Einleitung

(Multi)klassifikation

Binärisierung

Round Robin

Effiziente Klassifizierung

Dekodierung

Quick Weighted Voting

Effizientes Ranking

Ranking

Dekodierung

„Schweizer-System“

Ausblick

weitere Ansätze

maschinelle Klassifikation

Beschreibung

Ein Vorgang oder eine Methode zur automatischen Einteilung von Objekten in Klassen oder Kategorien

Gegeben:

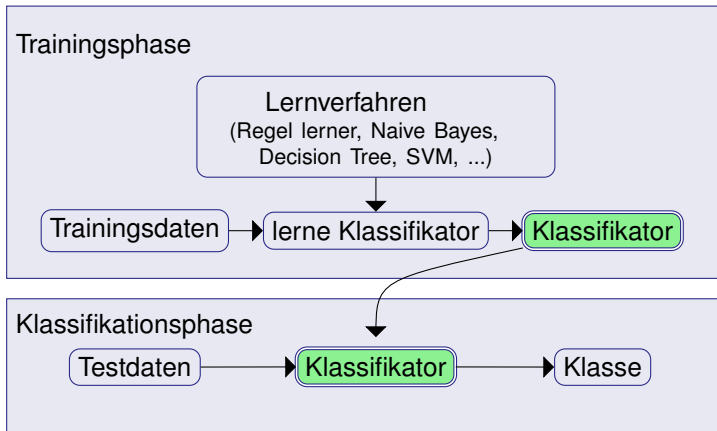
- ▶ Menge von Klassen $K = \{K_1, K_2, K_3, \dots, K_n\}$
- ▶ Menge von Daten $T = \{t_1, t_2, t_3, \dots, t_p\}$
- ▶ Trainingsdaten sind durch Attribut-Wert Beziehungen beschrieben, $t_i = \{w_1, w_2, w_3, \dots, w_m\}$, $w_j \in A_j$

Gesucht:

Eine Funktion $f : T \rightarrow K$ (der Klassifikator), die jedem Objekt die korrekte Klasse zuordnet.

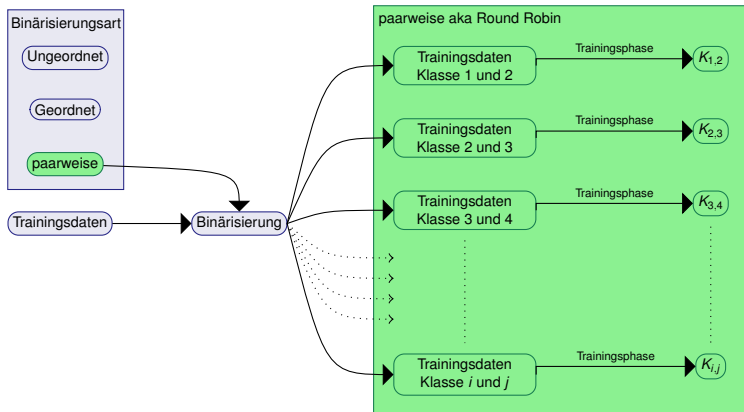


Schema der Klassifikation



Schema der Multiklassifikation

Trainingsphase mit Round Robin

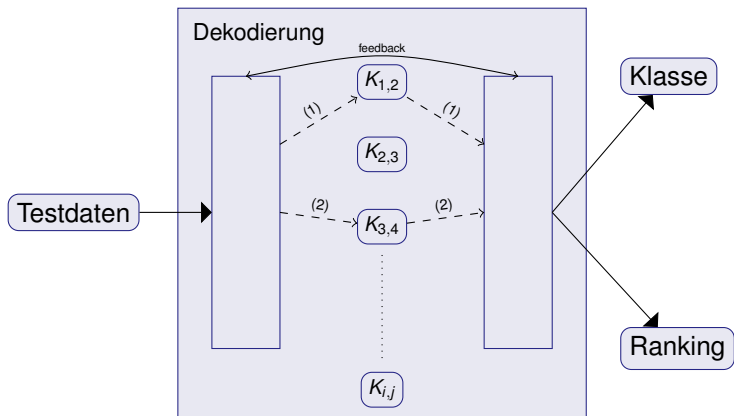


Bei der paarweisen Binärisierung entstehen $\frac{n(n-1)}{2}$ Klassifikatoren.



Schema der Multiklassifikation

Klassifikationsphase mit Round Robin



Round Robin Binärisierung [Fürnkranz, 2002]

Vorteile

- ▶ Genauigkeit
 - ▶ nie gegen *1-vs-rest* verloren
 - ▶ oft signifikant besser
- ▶ Effizienz
 - ▶ beweisbar schneller als *1-vs-rest*, ECOC, boosting,...
 - ▶ higher gains for slower base algorithms
- ▶ Einfachheit/Verständniss
 - ▶ einfachere Klassifikatoren → Hilfe für Konzeptlernen
- ▶ Datenreduktion
 - ▶ wesentlich geringere Trainingsmenge für jedes Binärproblem als für das ursprüngliche Problem
 - ▶ Teilprobleme passen komplett ins Speicher, wohingegen das ursprüngliche P. nicht
- ▶ Parallelisierbarkeit
 - ▶ unabhängige Teilprobleme



Round Robin Binärisierung

Nachteile

Bei n Klassen

- ▶ Trainingsphase: $\frac{n(n-1)}{2}$ Klassifikatoren zu lernen
- ▶ Klassifikationsphase : auszuwerten

Der Gesamtaufwand in der **Trainingsphase** ist trotzdem kleiner als n Lernprobleme in der *1-vs-rest* Binärisierung.

Diplomarbeitsziel

Effiziente Klassifikationsphase (Dekodierung) mit RR für Multiklassifikation und Ranking



Gliederung

Einleitung

(Multi)klassifikation

Binärisierung

Round Robin

Effiziente Klassifizierung

Dekodierung

Quick Weighted Voting

Effizientes Ranking

Ranking

Dekodierung

„Schweizer-System“

Ausblick

weitere Ansätze



Dekodierung

Votingbasiert

Voting

- ▶ Jeder Klassifikator $K_{i,j}$ stimmt für eine Klasse.
- ▶ Die Klasse mit den meisten Stimmen "gewinnt".

Weighted Voting

- ▶ Gewichtung der Stimmen
- ▶ Gewichtung z.B. durch Fehlerabschätzung der Vorhersage

Bei beiden Voting-Arten werden alle Klassifikatoren ($\frac{n(n-1)}{2}$) ausgewertet.

Dekodierung

andere Ansätze

Voting Variante: Vote Against [Cutzu, 2003]

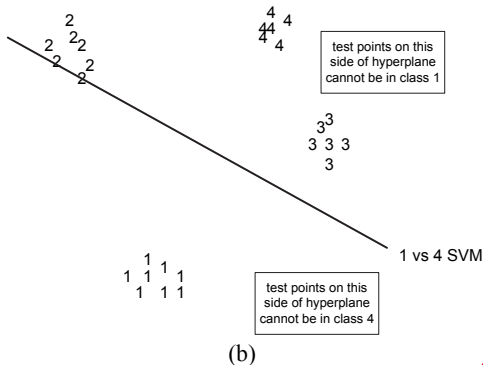
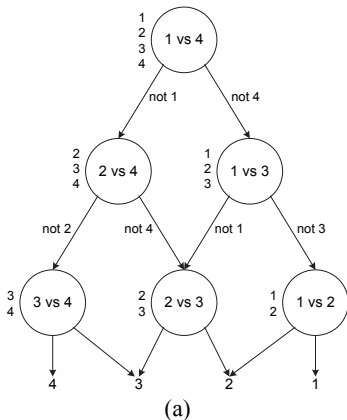
- ▶ Neue Interpretation für $k_1 \succ k_2$: Stimme **gegen** k_2 , anstatt **für** k_1
- ▶ bei Auswertung aller Vergleiche : Ergebnis identisch
- ▶ bei Auswertung einer Teilmenge der Vergleiche : Ergebnis besser

Decision Directed Acyclic Graph [Platt et al., 2000]

- ▶ Anordnung der Vergleiche im azyklisch gerichteten Graphen
 - ▶ Blätter des Graphen sind die vorhergesagten Klassen.
 - ▶ Pfad von Wurzel bis zum Blatt - vollständige Dekodierung
- ▶ nur $n - 1$ Vergleiche nötig

Dekodierung

andere Ansätze



quick Weighted Voting

Weighted Voting

- ▶ intuitiv einfach
- ▶ gute Genauigkeit
- ▶ Standarddekodierungsmethode
- ▶ Jedoch $\frac{n(n-1)}{2}$ Vergleiche nötig (auf den ersten Blick)

quick Weighted Voting

Beobachtung: Es gibt Vergleiche, die unnötig für die **Klassifikation** sind.

- ▶ Eigenschaft *uneinholbar*
 - ▶ Klasse j mit bisher $n - 2$ Gewinnen kann gegen Klasse i mit bisher 3 Verlusten nicht verlieren \rightarrow Klasse i kommt als *top rank* nicht mehr in Frage.
 - ▶ um oft diese Situation herzustellen \rightarrow *Reihenfolge* der Vergleiche wichtig
- ▶ andere Sichtweise (rate und verifiziere)
 - ▶ Wenn $\max_{i \in K}(\text{Voting}_i)$ und $\text{rank}_{\text{top}} = j$ gegeben, werden maximal $n - 1$ Vergleiche für die Verifikation benötigt.
 - ▶ Umformulierung als Suchproblem



quick Weighted Voting

Algorithmus

K sei die Menge der Klassen $K := \{k_1, k_2, k_3, \dots, k_n\}$ und
 $limit(x) := played(x) - wins(x)$.

1. Wähle Klasse k_i mit minimalen $limit(k_i)$ und $k_i \in K$
(Spieler₁)
2. Wähle Klasse k_j mit minimalen $limit(k_j)$, $k_j \in K \setminus k_i$ und die
Paarung (k_i, k_j) wurde noch nicht gespielt (Spieler₂)
3.
 - 3.1 Existiert keine Klasse, die die Bedingungen in Schritt 2
erfüllen \rightarrow winner := k_j STOP
 - 3.2 Ansonsten
 - 3.2.1 play(k_i, k_j)
 - 3.2.2 Aktualisiere Statistiken (limits, played, ...)
 - 3.2.3 Gehe zu Schritt 1

quick Weighted Voting

Vereinfacht

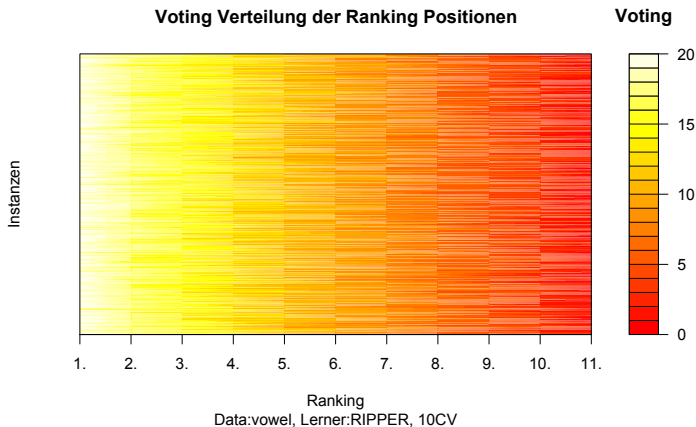
- ▶ Spiele solange ungespielte Paarungen mit der aktuellen top rank Klasse (bester limit bisher)
 - ▶ bis zur Verifikation vom top rank (limit korrekt)
 - ▶ Eventuell wird nach einem Spiel eine neue top rank Klasse bestimmt.

Zum Algorithmus

- ▶ Reihenfolge wichtig → Heuristik mit Limit
- ▶ Neben Reihenfolge, basiert die Effizienz auf eine Annahme über die Ergebnisse des Weighted Voting
- ▶ ein worst-case: alle Vergleiche liefern einen „Gleichstand“
 - ▶ Ergebnis ist eine gleichmässige Verteilung der Stimmen
 - ▶ kein Vergleich kann ausgelassen werden
 - ▶ alle Klassen potentieller top rank

quick Weighted Voting

(Weighted) Voting Verteilungen



quick Weighted Voting

Ergebnisse

Datensatz	n	$\frac{n(n-1)}{2}$	QWeighted
vehicle	4	6	3,9844
glass	6	15	9,7476
image	8	28	8,7459
yeast	10	45	15,8749
vowel	11	55	17,4212
soybean	19	171	27,6510
letter	26	325	45,0053

Tabelle: n ist die Anzahl der Klassen, $\frac{n(n-1)}{2}$ die Anzahl der Vergleiche von Weighted Voting und QWeighted die durchsch. Anzahl der Vergleiche von Quick Weighted Voting (Ripper, 10CV)



quick Weighted Voting

Ergebnisse

Datensatz	JRip	NaiveBayes	C4.5(J48)	SVM	
vehicle	3,9844	4,2660	3,9581	3,6442	6
glass	9,7476	9,5764	9,6896	9,9249	15
image	8,7459	9,0299	8,5472	8,2905	28
yeast	15,8749	15,8608	15,4774	15,5169	45
vowel	17,4212	17,0939	17,1253	15,2808	55
soybean	27,6510	27,6972	29,4460	28,3645	171
letter	45,0053	44,3960	47,7705	42,2618	325

Tabelle: Quick Weighted Voting unter verschiedenen Basis-Lerner



quick Weighted Voting

Eigenschaften

- ▶ Klassifikationsergebnis stimmt immer überein mit *Weighted Voting* (Bis auf mehrdeutige top ranks)
- ▶ Übernimmt somit die Eigenschaften (gute Genauigkeit, keine neuer empirischer Vergleich mit anderen Dekodierungsmethoden nötig)
- ▶ Reduzierung der Vergleiche nahezu unabhängig vom Basis-Lerner
- ▶ im worst-case Anzahl der Vergleiche identisch mit *Weighted Voting*
- ▶ im average-case $n \log(n)$ und best-case $n - 1$
- ▶ Ähnlichkeit zum A* Algorithmus



Gliederung

Einleitung

(Multi)klassifikation

Binärisierung

Round Robin

Effiziente Klassifizierung

Dekodierung

Quick Weighted Voting

Effizientes Ranking

Ranking

Dekodierung

„Schweizer-System“

Ausblick

weitere Ansätze



Ranking

Gegeben:

gleiche Ausgangsstellung wie bei Klassifikation

Gesucht:

Funktion, die jedem Testobjekt eine totale **Ordnung** der Klassen zuordnet

- ▶ Verallgemeinerung der Klassifikation

Dekodierung

- ▶ Votingbasiert
- ▶ SLATER-Optimal, Sortieralgorithmen (bei transitiven Vergleiche)
- ▶ Für Ranking grundsätzlich jede Klassifikationsdekodierung anwendbar (iterated choice [Hüllermeier et al., 2005]) mit:
 - ▶ bestimme top rank
 - ▶ entferne top rank und wiederhole mit restlichen Klassen

Ansatz der Diplomarbeit

Aufgrund der Analogie, Turniersysteme aus dem Sport, speziell das „Schweizer-System“ beim Schach

Schweizer-System [FIDE]

- ▶ ist eine Turnierform, die bei Schach- und ähnlichen Sportarten benutzt wird
- ▶ entstand aus der Notwendigkeit, dass bei hoher Spieleranzahl unmöglich jeder gegen jeden spielen kann
- ▶ Idee:
 - ▶ Es spielen immer soweit möglich gleichstarke Spieler in jeder Runde. ($\lfloor \frac{n}{2} \rfloor$ Paarungen)
 - ▶ Rundenzahl ist begrenzt. (z.B. $\log(n)$)
- ▶ Dadurch Anzahl der Spiele drastisch reduziert
- ▶ Spiel wird interessant, da fast jedes Spiel ein direkter Platzierungskampf bedeutet
- ▶ Anders als bei Systemen mit KO-Runden komplettes Ranking möglich

Maßstab [Hüllermeier et al., 2005]

verschiedene Anwendungsgebiete der Rankings -
verschiedene Semantik, mehrere Fehlerarten

Ranking Error

Position **jeder** Klasse
relevant

Differenz zwischen zwei
Rankings τ und $\tilde{\tau}$:

$$D_R(\tau, \tilde{\tau}) := \sum_{i=1}^m (\tilde{\tau}(i) - \tau(i))^2$$

Position Error

Position **einer** Klasse k_l
relevant

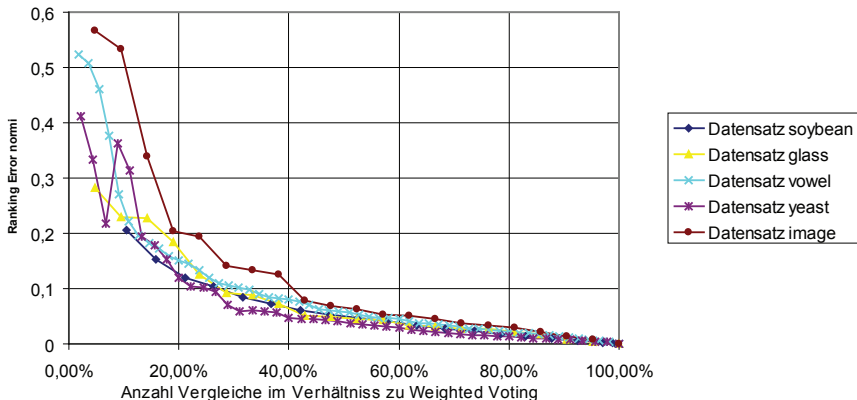
Differenz zwischen zwei
Rankings τ und $\tilde{\tau}$:

$$D_P(\tau, \tilde{\tau}) := |\tilde{\tau}(k_l) - \tau(k_l)|$$



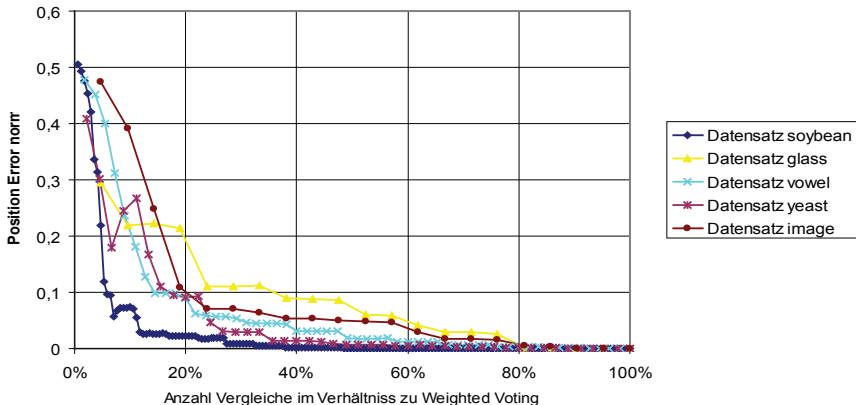
Ergebnisse

Ranking Error zwischen Swiss System und Weighted Voting



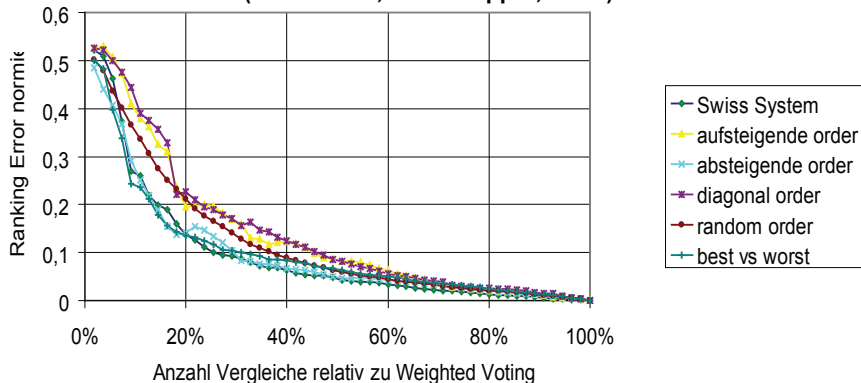
Ergebnisse

Position Error zwischen Swiss System und Weighted Voting



Ergebnisse

**Ranking Error mit verschiedenen Reihenfolgen
(Data:Vowel,Lerner:Ripper, 10CV)**



Ergebnisse

- ▶ Ranking-Error
 - ▶ Schweizer System immer besser oder gleich zufälliger Reihenfolge
 - ▶ Vermutlich alle Vergleiche nötig für korrektes Ranking
 - ▶ Reduzierung der Vergleichszahl ohne akzeptabler Verlust an Genauigkeit fraglich
- ▶ Position-Error
 - ▶ nicht alle Vergleiche nötig für top rank Bestimmung
 - ▶ Je mehr Klassen desto größer die prozentuale Reduzierung der nötigen Vergleiche
 - ▶ obige Beobachtungen im Einklang mit QWeighted



Fazit

- ▶ Eine geschickte Reihenfolge verringert schneller den Ranking-Error
- ▶ Jedoch mit Schweizer-System allein keine akzeptable Reduzierung der Vergleiche möglich
- ▶ Zusätzlich zum Schweizer-System Votingmethodik erweitern (→ weitere Ansätze), die evt. nicht mehr identisch mit Weighted Voting Ergebniss sind, jedoch eine bessere Genauigkeit nach weniger Vergleichen liefert.



Gliederung

Einleitung

(Multi)klassifikation

Binärisierung

Round Robin

Effiziente Klassifizierung

Dekodierung

Quick Weighted Voting

Effizientes Ranking

Ranking

Dekodierung

„Schweizer-System“

Ausblick

weitere Ansätze



Vererbung

- ▶ Annahme:

$$P(k_i >_{\tau} k_j | k_i \succ k_j) > 0.5$$

- ▶ also: die Vergleiche sind eher transitiv bzgl. des Ranking
- ▶ Algorithmus votingbasiert, zusätzlich zur gewichtete Stimme eine Art Bonus
- ▶ Bonus wird aus den bisherigen Stimmen von den verlorenen Gegnerklassen berechnet (Geerbt)

$$B_{k_j} = \sum_{k_i \in K, k_j > k_i} W_{k_i} * T(k_j, W_{k_j}, Played(k_j))$$

$$W_{k_j}^+ = W_{k_j} + B_{k_j}$$



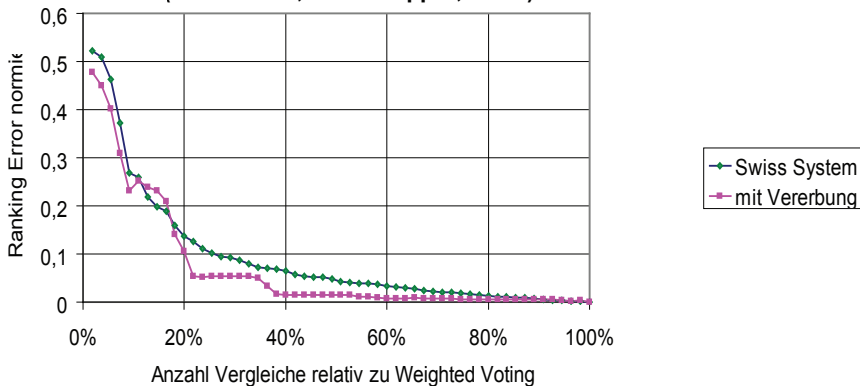
Vererbung 2

- ▶ sollte folgende Eigenschaften besitzen:
 - ▶ Je mehr Siege eine Klasse, desto höherer Bonus/ mehr geerbte Stimmen
 - ▶ Je mehr Spiele gespielt desto grösser der Transitivitätsfaktor
 - ▶ Jedoch Bonus/Transitivitätsfaktor so ausgelegt, dass keine falsche Dominanz entsteht
 - ▶ Eventuell bei voller Anzahl der Vergleiche, Ranking identisch mit dem Weighted Voting Ranking
- ▶ Sinnvoll in Situationen, in denen mehrere Klassen bisher gleiches Verhältnis zwischen Stimmen und Spielen haben
→ Rankingposition mehrdeutig



Vererbung 3

Ranking Error
(Data:Vowel,Lerner:Ripper, 10CV)



weitere Schritte

- ▶ formale Komplexitätsanalyse von Quick Weighted Voting
- ▶ Vererbungsmethode testen
- ▶ Laufzeitanalysen

Literaturverzeichnis



J. Fürnkranz.

Round robin classification.

Journal of Machine Learning Research, 2:721–747, 2002.



J. C. Platt, N. Cristianini and J. Shawe-Taylor.

Large margin DAGs for multiclass classification.

Advances in Neural Information Processing Systems 12 (NIPS-99), pages 547–553. MIT Press, 2000.



World Chess Federation - Federation Internationale des Echecs.

FIDE Swiss Rules

FIDE Handbook, chapter C.04

<http://www.fide.com/official/handbook.asp?level=C04>;



E. Hüllermeier, J. Fürnkranz and J. Beringer.

On Position Error and Label Ranking through Iterated Choice.

LWA, pages 158–163. DFKI, 2005.



F. Cutzu.

Polychotomous Classification With Pairwise Classifiers: a new Voting Principle.

Multiple Classifier Systems, 4th International Workshop, MCS 2003, pages 115–124. Springer, 2003.

