

Atıfların Otomatik Kategorizasyonu için Örnek Uygulama

Zehra Taşkın

ztaskin@hacettepe.edu.tr

Plan

- ❑ Atıf analizlerinde mevcut durum
- ❑ Makine öğrenmesi yöntemleri ve atıf analizlerine uygulanması
- ❑ Örnek uygulama
- ❑ Sonuç ve değerlendirme

Mevcut durum



zehrataskin

@zehrataskin

Takip et



World is changing.
Scientometrics must be, too.
This transformation is needed.
#STI2017

Summary: Towards transdisciplinary collaborations

Transdisciplinary teams needed for:

- Participation of relevant stakeholders
- Experts on the sector under study to use and interpret data
- STS – interviews, ethnography, participation
- Experts on 'technologies of participation' (Rip, Doug Robinson)

Complementarities between Scientometrics and STS
(gross simplification of traits)

Scientometrics	STS
Secluded research	Research in the wild
Positivist	Interpretative
Value-free	Value-laden
Technocratic State	Civil Society
Quantitative	Qualitative
Top-down	Bottom-up
Expert-based	Participatory
Closing down	Opening up

Back to the origins of scientometrics?

De Solla Price
Henry Small
rooted on
History of Science

Garfield's ISI close to
U. Penn Dept. of Sociology
and History

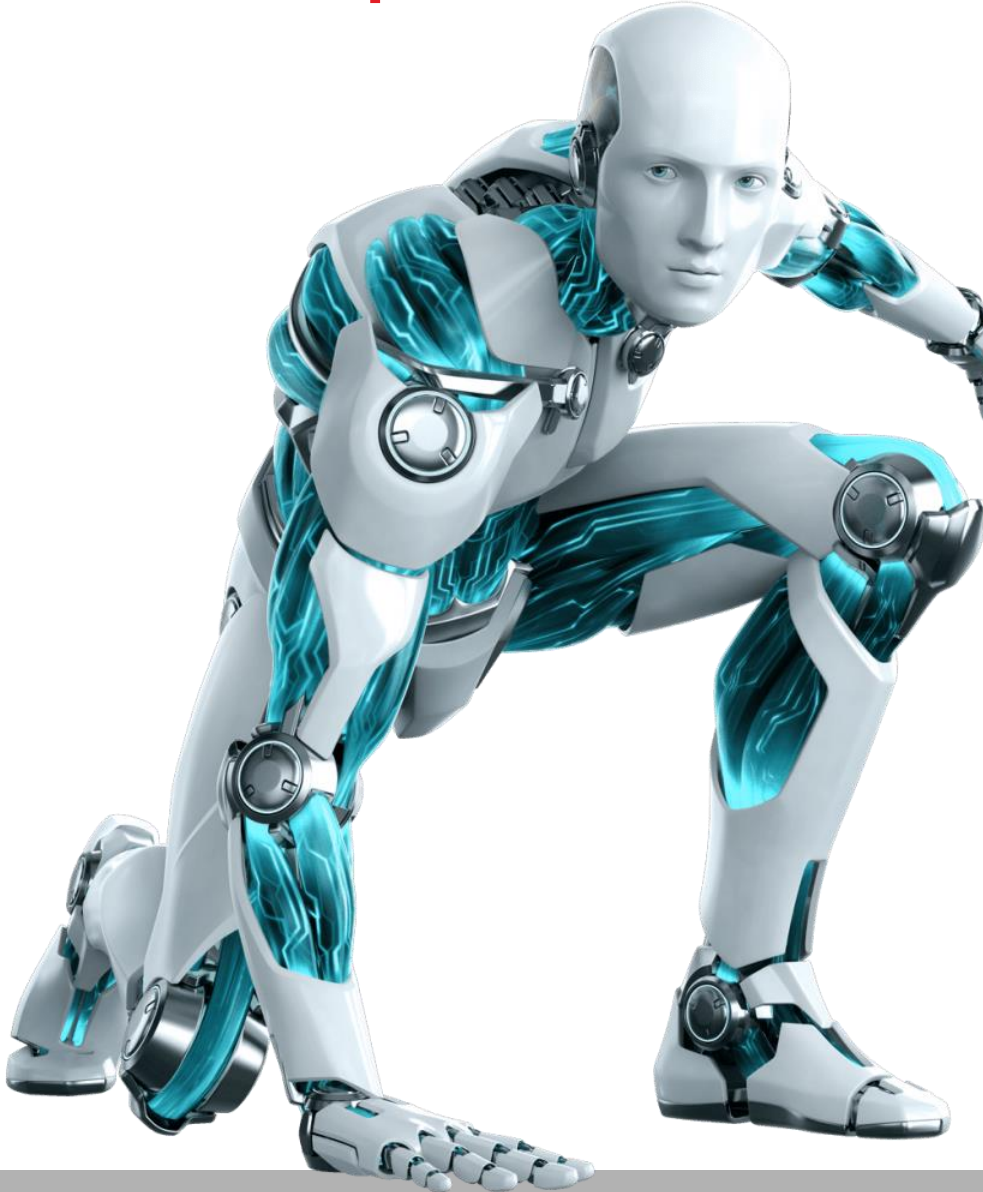
03:29 - 06 Eyl 2017

8 Retweet 2 Beğeni



- ❑ "ölümcül günah"
- ❑ Saymazsak nasıl değerlendireceğiz?
- ❑ Alternatif var mı?
- ❑ Saymanın yerini içerik analizi alabilir mi?
- ❑ Süper kahraman kim?

Süper kahraman: Makine



- ❑ Eğitilebilir
- ❑ Öğrenebilir
- ❑ Öğrendiklerini benzer verilere uygulayabilir
- ❑ Atıflar için ne yapabilir?

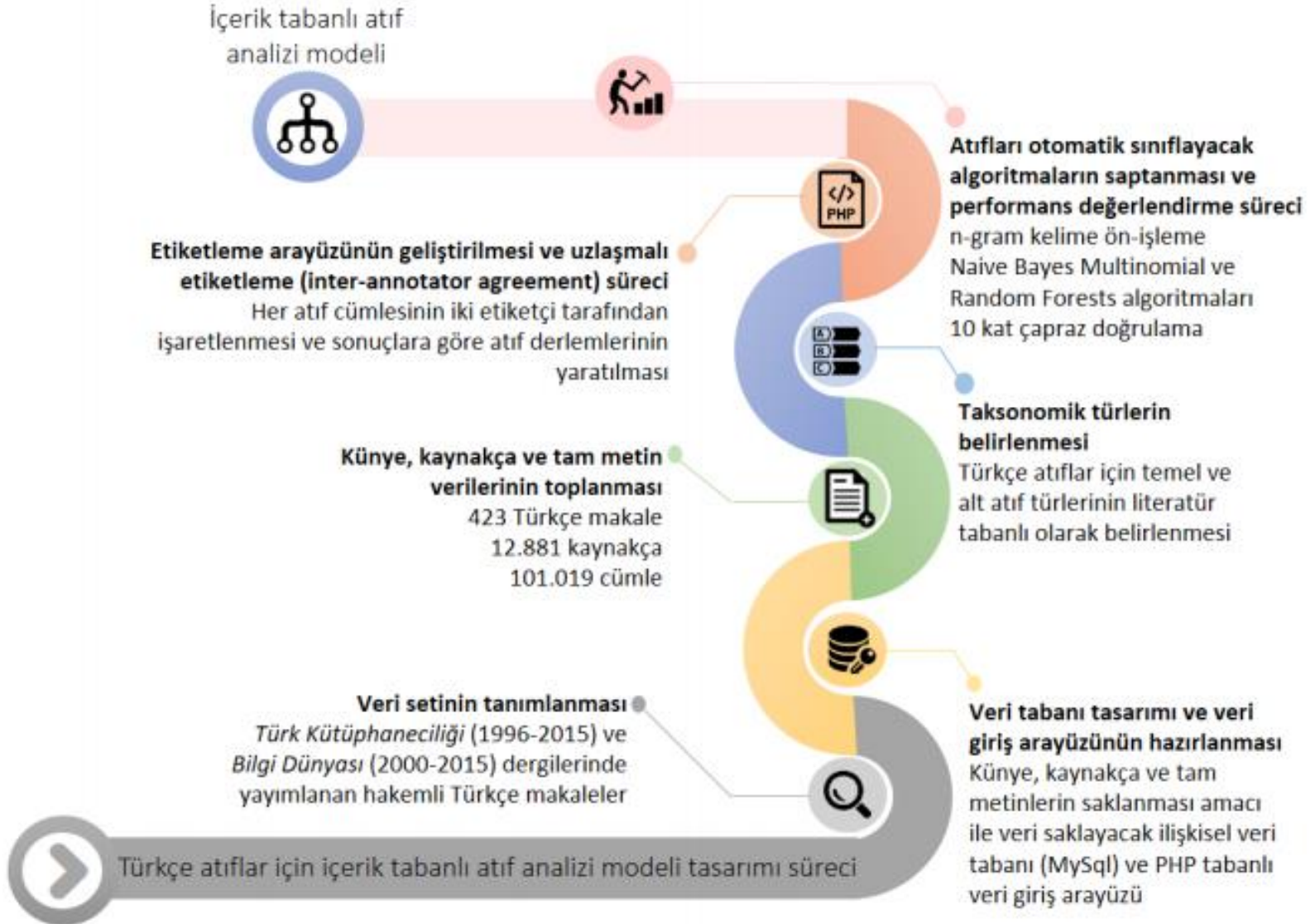
Yapabilecekleri...

- ❑ Atıfları metin içinden saptayabilir
 - ❑ Kaynakça?
 - ❑ Gönderme?
 - ❑ Atıf cümlesi?
- ❑ Saptadığı atıf cümlelerini sınıflandırabilir
 - ❑ Nasıl bir sınıflandırma?
- ❑ Sınıflanmış veriyi yorumlayabilir

Eđitcinin eđitimi?

- Makineyi eđitmek gerekli
 - Nasıl?
- 115K440 - Trke atıflar iin ierik tabanlı atıf analizi modeli tasarımı
 - Ama: Makineyi ierik tabanlı atıf analizi grevlerini gerekleřtirmesi iin eđitmek

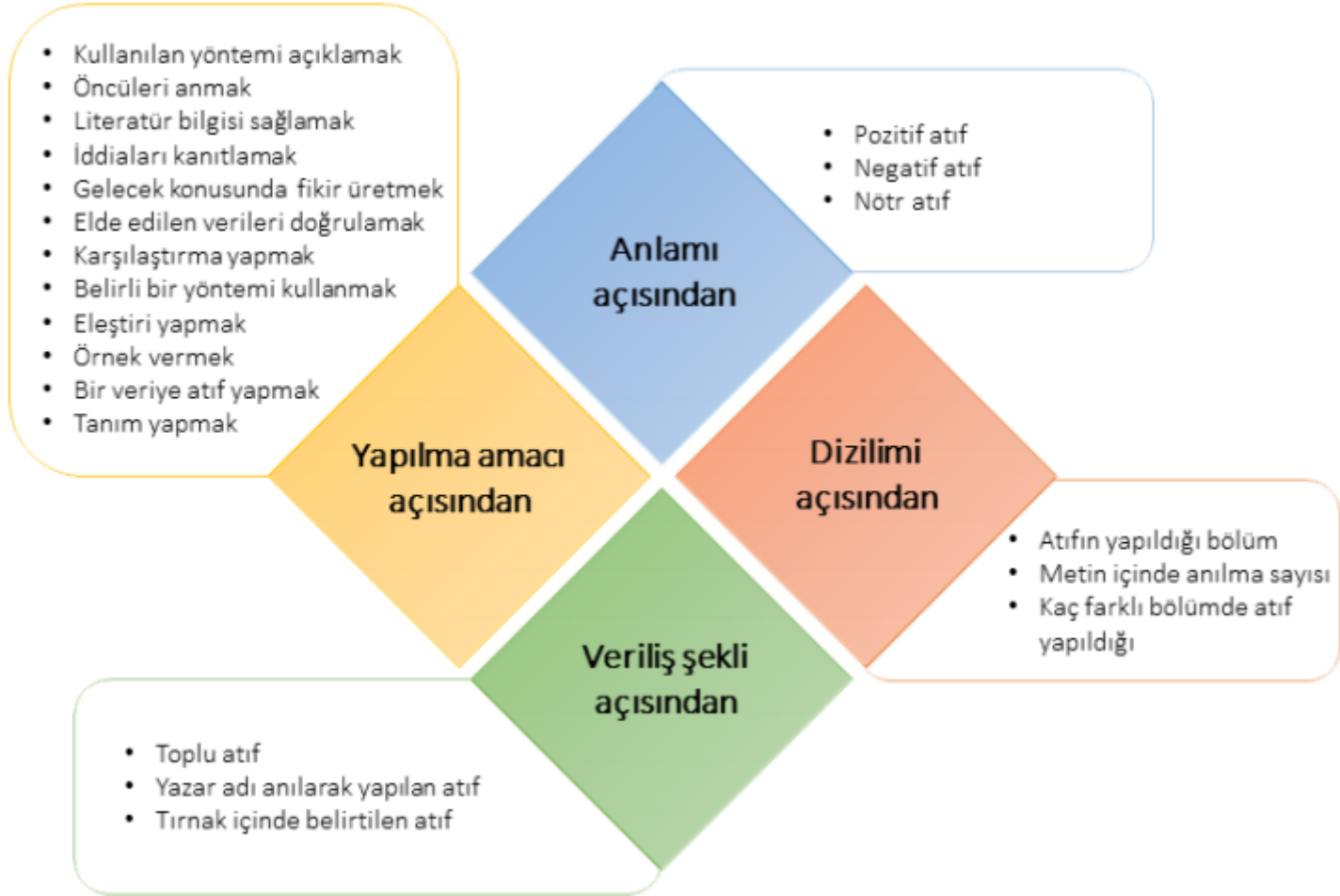
Süreç akışı



Proje ekibi

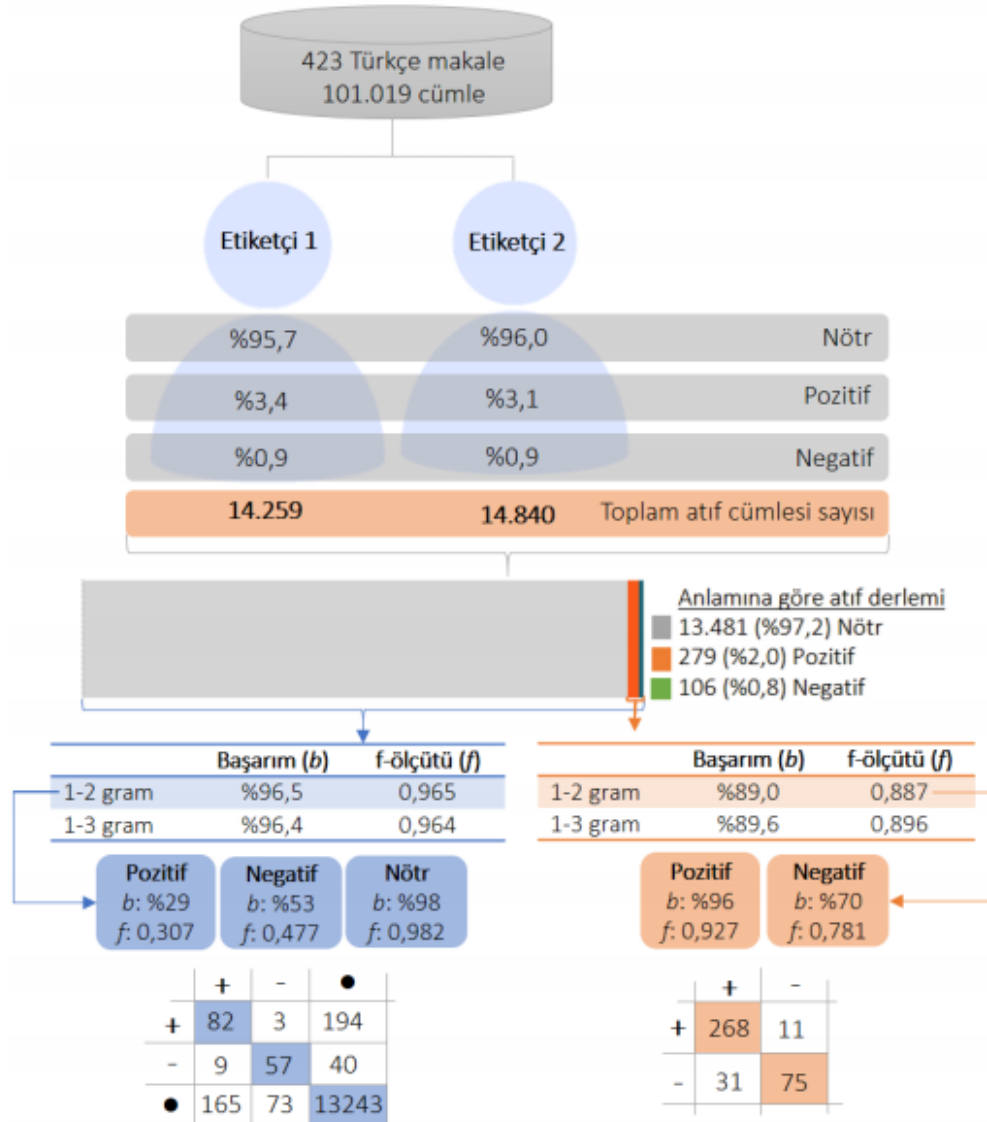
- ❑ 6 uzman
 - ❑ 4 doktora, 2 lisans öğrencisi
- ❑ Uzlaşmalı etiketleme süreci
 - ❑ Neden?
 - ❑ Ortak karar
 - ❑ Atıf cümlesinin sınıfına karar vermek kolay değil

Taksonomik atıf sınıfları

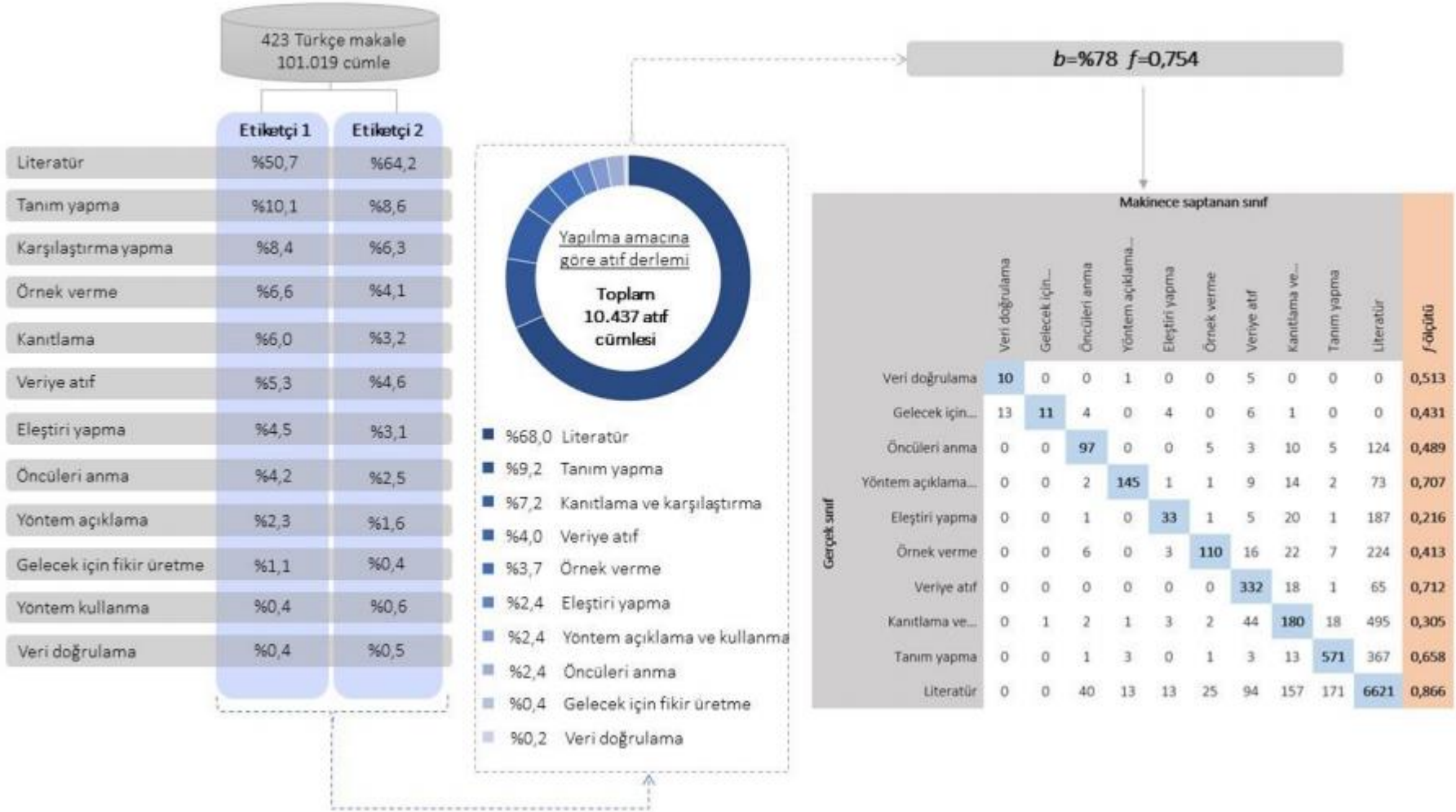


□ http://115k440.hacettepe.edu.tr/etiketleme_egitimi.jpg

Anlamı açısından atıflar



Amacı açısından atıflar



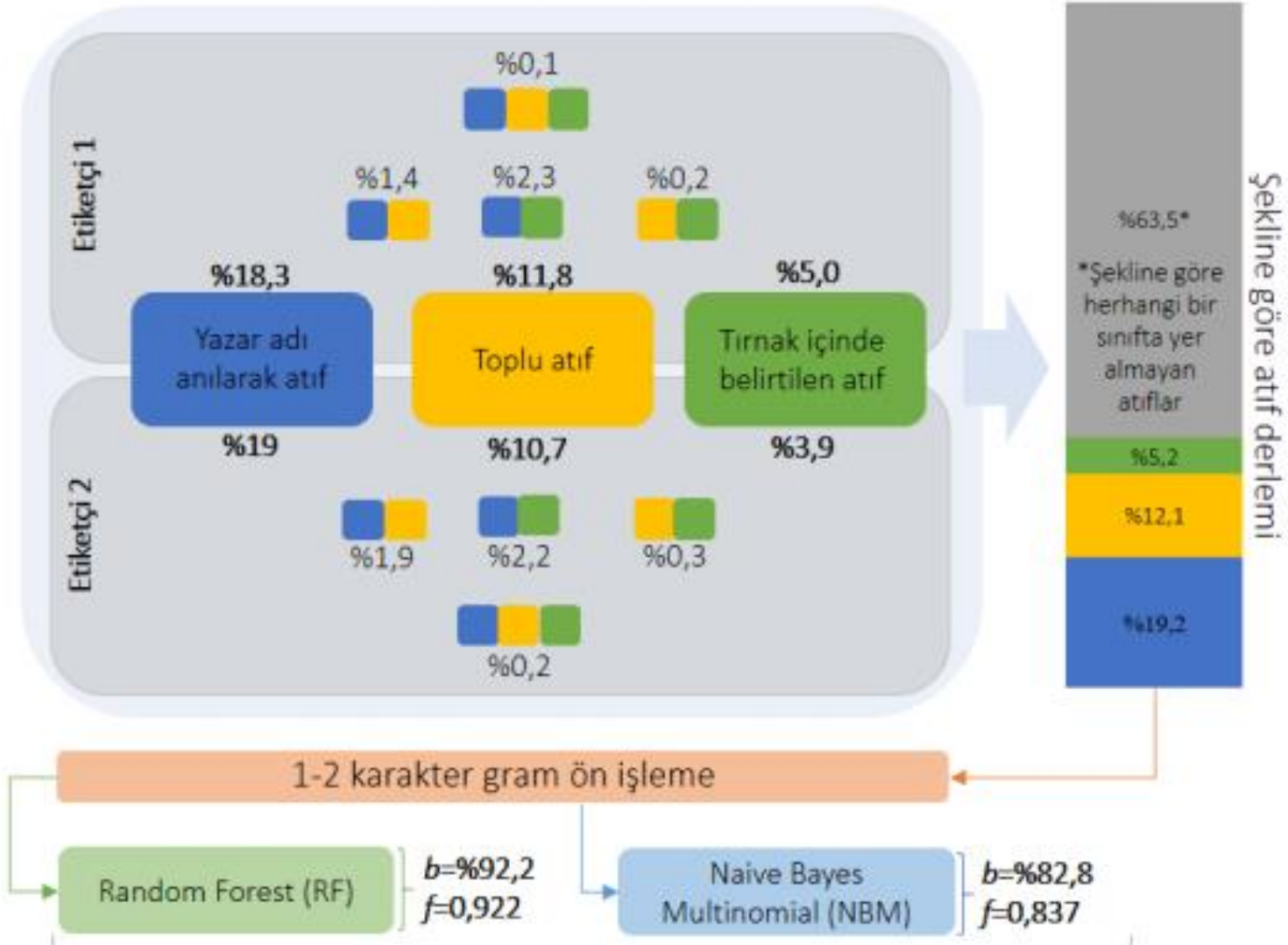
Amacı açısından atıflar

%90,4

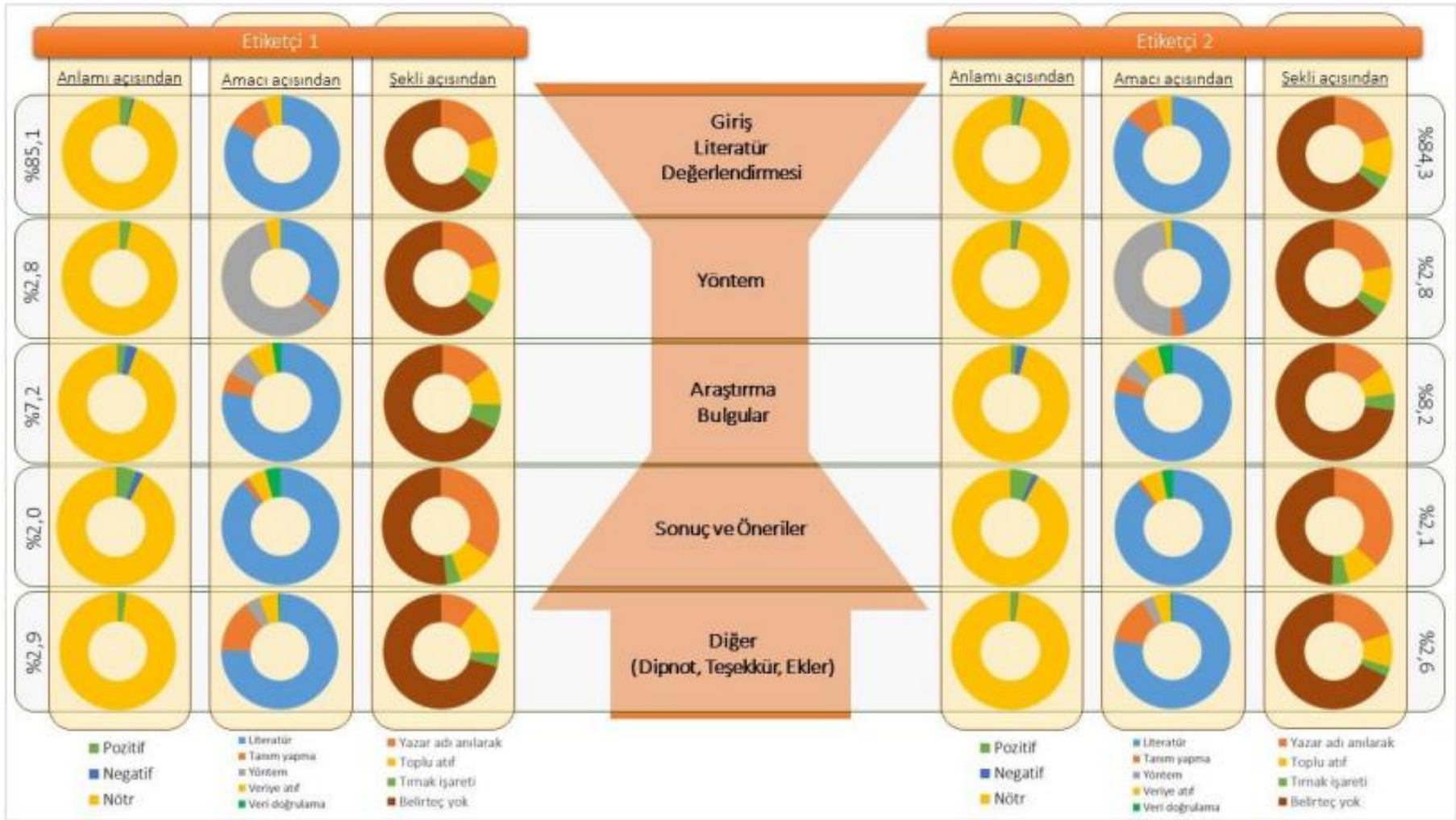
Makinece saptanan sınıf

	Literatür	Tanım yapma	Yöntem	Veriye atıf	Veri doğrulama	Başarım	f-ölçütü	
Gerçek sınıf	Literatür	8286	270	30	207	6	%94	0,944
	Tanım yapma	328	621	4	6	0	%65	0,670
	Yöntem	73	3	161	10	0	%65	0,729
	Veriye atıf	61	0	0	355	0	%85	0,714
	Veri doğrulama	0	0	0	0	16	%100	0,842

Şekli açısından atıflar



Dizilimi açısından atıflar



Kaynakçada yer alan künyelerin %6'sına metin içinde gönderme yok!
Metin içinde yer alan göndermelerin %1'i kaynakçada yok!

Sonuç

- ❑ Tüm sınıflarda en az %90 başarı
- ❑ Kural listeleri
- ❑ Büyük yoğunluklu metinlerden atıf cümlelerinin çıkarımı
 - ❑ Başarının tanımı: Az insan gücü ile yüksek başarı oranları

Çıkarım sistemi



SCIENCE, TECHNOLOGY & INNOVATION INDICATORS

Open indicators: innovation, participation and actor-based STI indicators
PARIS 2017

First Stage of an Automated Content-Based Citation Analysis Study: Detection of Citation Sentences¹

Zehra Taşkın^{*}, Umut Al^{*} and Umut Sezen^{**}

Atıf analizlerinin geleceđi



Illustration: David Parkins => <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>

Atıfların Otomatik Kategorizasyonu için Örnek Uygulama

Zehra Taşkın

ztaskin@hacettepe.edu.tr