

21-055 vom 08.06.2021

Statistische Berechnung

## EM-Prognose – Frankreich wird Fußball-Europameister

Nach dem Weltmeistertitel könnte auch jener der Fußball-EM in Frankreich landen – zu diesem Ergebnis kommen Forscherinnen und Forscher der Universitäten Innsbruck und Gent, der Technischen Universitäten Dortmund und München und der Hochschule Molde. Gute Chancen auf den Titel dürfen sich aber auch England und Spanien ausrechnen. Die Berechnung ist auch Thema beim nächsten Vortrag der Reihe „Zwischen Brötchen und Borussia“ der TU Dortmund am Samstag, 12. Juni.

Am Freitag, 11. Juni, starten Europas Herren-Fußballmannschaften mit einem Jahr Verspätung in die Europameisterschaft. Der Favorit ist diesmal Frankreich mit einer Gewinnwahrscheinlichkeit von 14,8 Prozent. Das zeigt ein internationales Forschungsteam bestehend aus Andreas Groll und Franziska Popp (beide TU Dortmund), Gunther Schauburger (TU München), Christophe Ley und Hans Van Eetvelde (beide Universität Gent), Achim Zeileis (Universität Innsbruck) und Lars Hvattum (Hochschule Molde, Norwegen) mit Hilfe von maschinellem Lernen. Ihre Prognose kombiniert dabei mehrere statistische Modelle für die Spielstärken der Teams mit Informationen über die Team-Struktur (etwa Marktwert, Anzahl Champions-League-Spieler, Vereinsspiele-Performance einzelner Spieler) sowie sozio-ökonomische Faktoren des Herkunftslandes (Bevölkerung und Bruttoinlandsprodukt).

Mit den vorhergesagten Werten aus dem Modell wurde die gesamte EM 100.000 Mal durchsimuliert: Spiel für Spiel, der Turnierauslosung und allen UEFA-Regeln folgend. Damit ergeben sich Wahrscheinlichkeiten für das Weiterkommen aller Teams in die einzelnen Turnierrunden und letztendlich für den EM-Sieg. Favorit ist Frankreich mit einer Gewinnwahrscheinlichkeit von 14,8 Prozent, gefolgt von England (13,5) und Spanien (12,3). Das Turnier ist natürlich dennoch nicht gelaufen – das zeigen auch die relativ knappen Abstände bei den Gewinnwahrscheinlichkeiten an der Spitze, außerdem die ohnehin niedrige Wahrscheinlichkeit selbst der Top-Nationen. „Es liegt in der Natur von Prognosen, dass sie auch danebenliegen können – sonst wären Fußball-Turniere auch sehr langweilig. Wir liefern eben Wahrscheinlichkeiten, keine Gewissheiten, und eine Gewinnwahrscheinlichkeit von 15 Prozent heißt zugleich, dass die Mannschaft zu 85 Prozent nicht Turniersieger werden kann“, sagt Achim Zeileis. Bisher waren die Prognosen aber durchaus erfolgreich: Das Innsbrucker Modell von Achim Zeileis, das auf bereinigten Quoten der Wettanbieter basiert, konnte unter anderem bereits 2008 das EURO-Finale sowie 2010 und 2012 Welt- und Europameister Spanien richtig vorhersehen. Dieses Jahr wird es als Teil des umfassenderen kombinierten Modells eingesetzt, das von den Teams um Andreas Groll (TU Dortmund), Gunther Schauburger (TU München) und Christophe Ley (Universität Gent) entwickelt wurde und das bei der Fußball-Weltmeisterschaft 2018 die Prognosegüte der Wettanbieter übertroffen hatte.

Die deutsche Nationalmannschaft wurde dieses Jahr in eine besonders herausfordernde Gruppe gelost: „In Gruppe F sind drei sehr starke Teams, darunter der amtierende Weltmeister Frankreich und der Europameister Portugal, beide zugleich die Finalisten der EURO 2016, plus eben Deutschland“, erläutert Andreas Groll: „In dieser Gruppe ist die Wahrscheinlichkeit deshalb verglichen mit den Favoriten in den anderen Gruppen geringer, es bis ins Achtelfinale zu schaffen. Wer das aber schafft, hat dann ganz gute Chancen, weiterzukommen.“ Die Prognose sieht sowohl für Deutschland als auch Portugal eine Wahrscheinlichkeit von 85,3 Prozent, es ins Achtelfinale zu schaffen; für Frankreich liegt diese Wahrscheinlichkeit etwas höher bei 89,7 Prozent. Dass Deutschland Europameister wird, ist mit 10,1 Prozent Wahrscheinlichkeit deutlich unter den Werten der Favoriten und genau gleichauf mit Portugal.

Die Berechnung der Forscherinnen und Forscher basiert auf vier Informationsquellen: Einem statistischen Modell für die Spielstärke jedes Teams auf Basis aller Länderspiele der vergangenen acht Jahre, einem weiteren statistischen Modell für die Spielstärke der Teams auf Basis der Wettquoten von 19 internationalen Buchmachern, zusätzlichen Informationen über die Teams – zum Beispiel der Marktwert – und über ihre Herkunftsländer – etwa die Bevölkerungszahl oder das Bruttoinlandsprodukt, außerdem detaillierten Ratings der einzelnen Spieler und deren individueller Performance. Ein Machine-Learning-Modell führte die vier Quellen zusammen und optimierte sie schrittweise. Die Forschenden haben das Modell zuvor mit historischen Daten trainiert, wie Andreas Groll erläutert: „Wir haben das Modell mit den jeweils zu dem Zeitpunkt aktuellen Daten für die vergangenen vier Europameisterschaften, also zwischen 2004 und 2016, gefüttert und mit den tatsächlichen Spielausgängen aller Spiele der jeweiligen Turniere vergleichen lassen – so wird die Gewichtung der einzelnen Informationsquellen für das aktuelle Turnier im Idealfall sehr genau ausfallen.“ Wie gut das Modell abgeschnitten hat, werden Fußballfans auf alle Fälle spätestens am Abend des 11. Juli erfahren. Wer schon vorher tiefer in das Thema einsteigen möchte, kann am Samstag, 12. Juni, um 10:45 Uhr an der beliebten öffentlichen Reihe „Zwischen Brötchen und Borussia – moderne Physik für alle“ der TU Dortmund teilnehmen: Prof. Andreas Groll erklärt in seinem Online-Vortrag die statistische Berechnung des Europameisters.

#### **Weitere Informationen:**

Die gesamte Prognose mit interaktiven Grafiken gibt es hier: <http://bit.ly/forecast-euro2020>

Alle Infos zu „Zwischen Brötchen und Borussia“:

<https://physik.tu-dortmund.de/oeffentlichkeit/zwischen-broetchen-und-borussia/>

**Bildhinweis:** Prof. Andreas Groll leitet an der Fakultät Statistik der TU Dortmund das Fachgebiet Statistical Methods for Big Data. Foto: Roland Baege/TU Dortmund

**Ansprechpartner für Rückfragen:**

Prof. Andreas Groll

Fakultät Statistik der TU Dortmund

Tel.: 0231 755-4229

E-Mail: [andreas.groll@tu-dortmund.de](mailto:andreas.groll@tu-dortmund.de)

Die Technische Universität Dortmund hat seit ihrer Gründung vor 52 Jahren ein besonderes Profil gewonnen, mit 17 Fakultäten in Natur- und Ingenieurwissenschaften, Gesellschafts- und Kulturwissenschaften. Die Universität zählt rund 33.400 Studierende und 6.500 Mitarbeiterinnen und Mitarbeiter, darunter etwa 300 Professorinnen und Professoren. Das Lehrangebot umfasst rund 80 Studiengänge. In der Forschung ist die TU Dortmund in vier Profildbereichen besonders stark aufgestellt: (1) Material, Produktionstechnologie und Logistik, (2) Chemische Biologie, Wirkstoffe und Verfahrenstechnik, (3) Datenanalyse, Modellbildung und Simulation sowie (4) Bildung, Schule und Inklusion. Aufgrund ihrer vorbildlichen Transferstrategie wird die TU Dortmund im „Gründungsradar 2020“ in der Spitzengruppe der großen Hochschulen gelistet. Bis zu ihrem 50. Geburtstag belegte die TU Dortmund beim QS-Ranking „Top 50 under 50“ Rang drei der bundesdeutschen Neugründungen.