



Predicting VEGF Levels in Cancer Patients Using Machine and Deep Learning

Theofili I.¹, Liolis E.², Karalis V.D.¹, Sivolapenko G.³

¹Department of Pharmacy, National and Kapodistrian University of Athens, Greece,

²School of Medicine, University of Patra, Patra, Greece,

³Department of Pharmacy, University of Patra, Patra, Greece



Introduction

Increased VEGF expression in the tumor microenvironment contributes to tumor progression and metastasis [1]. Bevacizumab, a recombinant humanized monoclonal antibody, is used as monotherapy or in combination with chemotherapy for the treatment of various types of cancers including breast cancer, cervical, ovarian cancer, non-small cell lung cancer (NSCLC) and metastatic colorectal (CRC) [2]. It suppresses tumor angiogenesis by preventing VEGF-A from interacting with its receptors, thus, inhibiting tumor growth [3]. Bevacizumab has shown remarkable versatility across a plethora of cancers, showing efficacy in frontline treatment, maintenance therapy, and recurrent disease settings [4].

Objectives

The aim of this study was to use machine learning algorithms to investigate the impact of demographic and chemotherapy factors as well as bevacizumab levels (peak, trough, and their difference) on VEGF levels in patients with NSCLC, CRC, breast, cervical and ovarian cancer.

Methods

Several machine learning techniques and Artificial Neural Networks (ANN) were used for the analysis. Initially, methods such as Categorical Principal Component Analysis (CATPCA), Multiple Correspondence Analysis (MCA), Principal Component Analysis (PCA) and Factor Analysis of Mixed Data (FAMD) were applied to explore patterns and variable associations. To model the relationship between the variables and the reduction in VEGF levels (i.e., VEGF_D) models such as Random Forest, Boosted Trees, Bagged Trees, and ANN were developed and evaluated.

Results

Both CATPCA and PCA revealed the opposing trends between VEGF and bevacizumab levels, suggesting different contributions to variance (Figure 1). CATPCA also indicated that VEGF reduction may follow distinct trajectories based on treatment duration and cycles, highlighting individualized responses (Figure 2 and 3). The random forest analysis identified AV_D, weight group, and number of chemotherapy cycles as key predictors of VEGF reduction, with AV_D having the strongest dose-dependent effect. Boosted trees Bagged trees and ANN confirmed previous findings identifying AV_D as the most decisive factor, while weight and chemotherapy duration remained significant but less influential. Overall, AV_D consistently exhibits the strongest effect on VEGF reduction (VEGF_D) across all models, with weight and treatment duration also playing an important role (Figure 4).

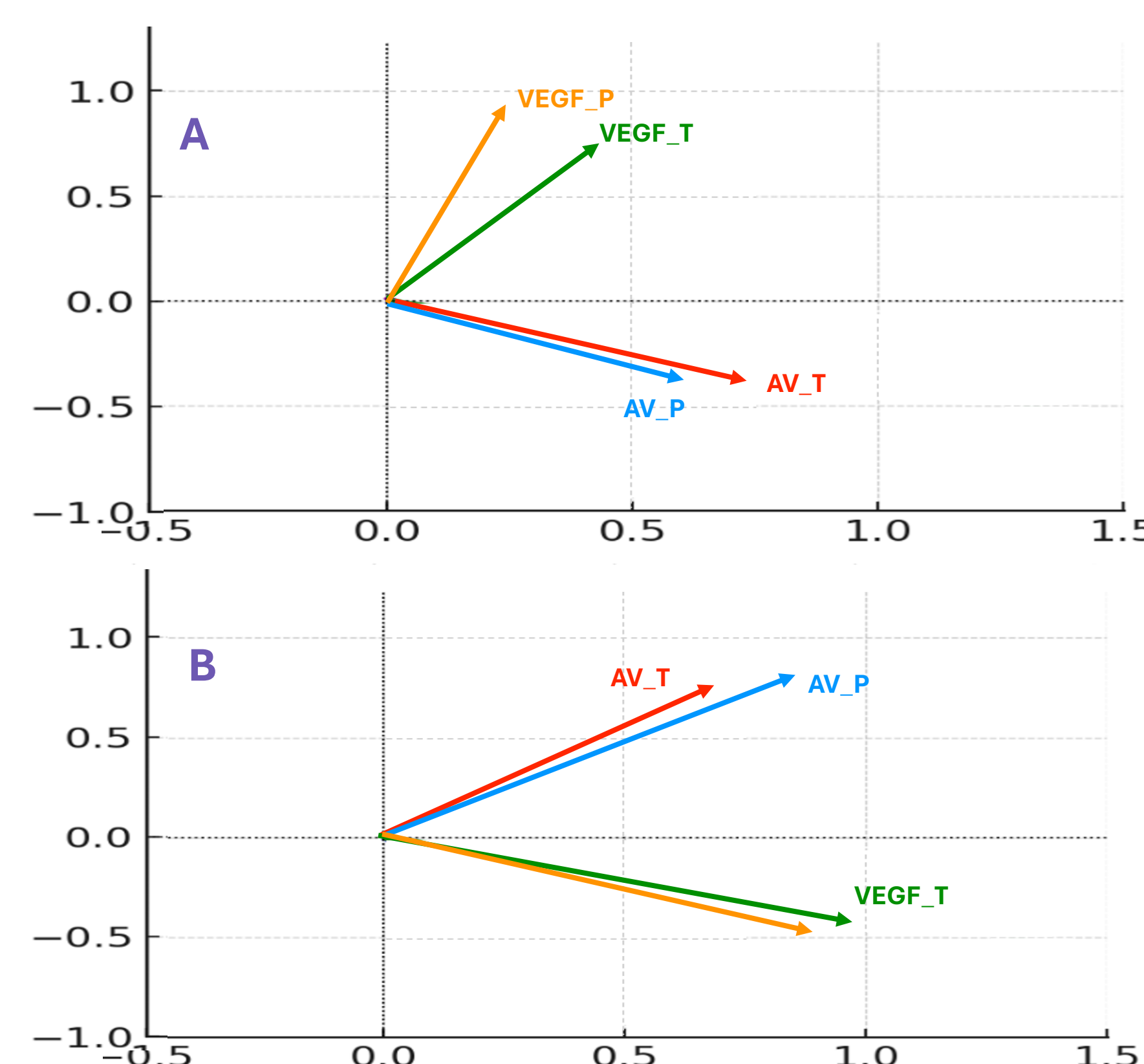


Figure 1. A. Categorical Principal component analysis (CAT-PCA) for the variables AV_P, AV_T, VEGF_P, and VEGF_T. B. Principal component analysis (PCA) for the variables AV_P, AV_T, VEGF_P, and VEGF_T.

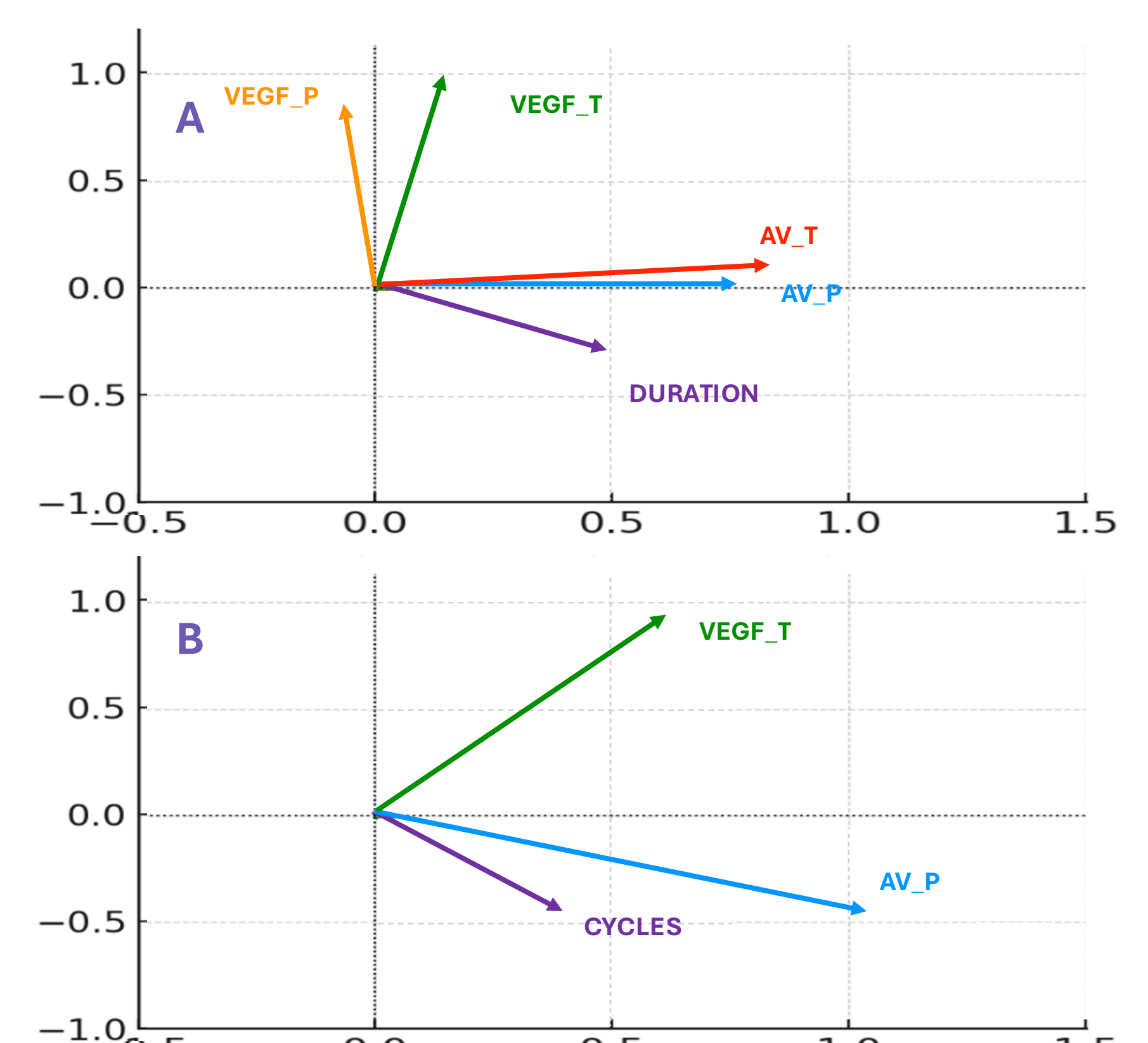


Figure 2. A. Principal component analysis (PCA) for the variables AV_P, AV_T, VEGF_P, VEGF_T, and Therapy duration. B. Categorical Principal component analysis (CAT-PCA) for the variables AV_P, AV_T, VEGF_P, VEGF_T, and Cycles of the therapy.

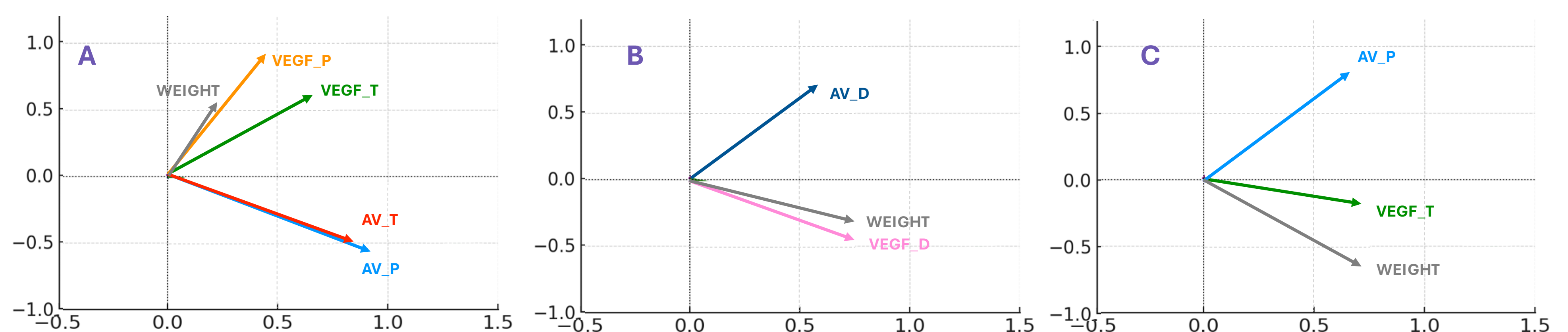


Figure 3. A. Principal component analysis (PCA) for the variables AV_P, AV_T, VEGF_P, and VEGF_T, and Weight. B. Principal component analysis (PCA) for the variables AV_D, VEGF_D, and Weight. C. Principal component analysis (PCA) for the variables AV_P, VEGF_T, and Weight.

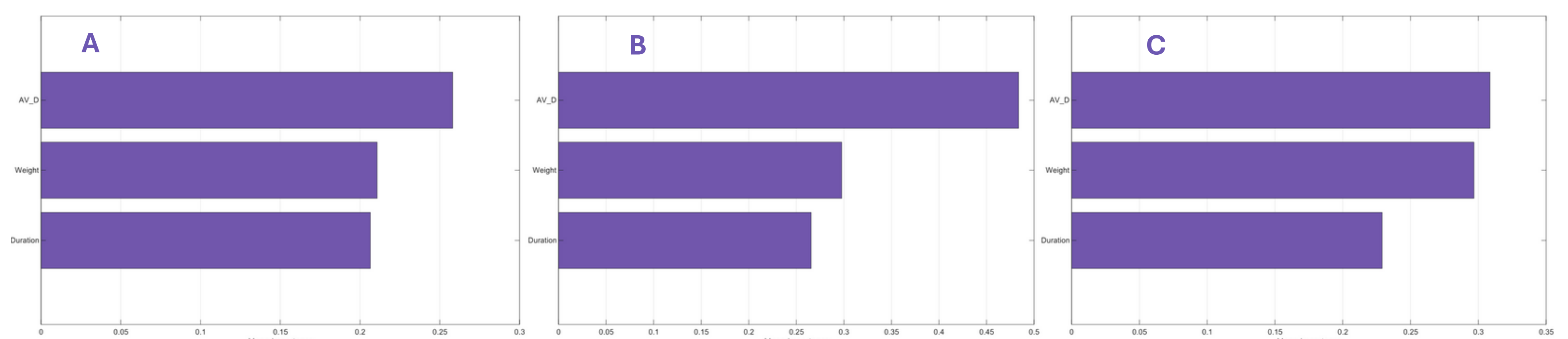


Figure 4. A. Boosted trees variable importance score for the variables AV_D, Weight, and Therapy duration. B. Bagged trees variable importance score for the variables AV_D, Weight, and Therapy duration. C. ANN variable importance score for the variables AV_D, Weight, and Therapy duration.

Conclusions

This study underlines the central role of increasing bevacizumab levels in reducing VEGF levels. However, it is not the absolute bevacizumab levels that matter, but rather the increase in those levels. The influence of weight and treatment duration suggests that individualized factors should be considered when optimizing therapy. This study clearly showed that machine learning techniques can uncover critical relationships between demographic characteristics and VEGF reduction, providing valuable insights for personalized treatment.

References:

1. Alidzanovic, L. et al. (2016). Oncotarget, 7(35), 57197–57212.
2. Han, K. et al. (2016). Cancer chemotherapy and pharmacology, 78(2), 341–351.
3. European Medicines Agency. (2023). Avastin EPAR product information. Retrieved from https://www.ema.europa.eu/en/documents/product-information/avastin-epar-product-information_en.pdf
4. Ghezelayagh, T. S. et al. (2023). European journal of gynaecological oncology, 44(1), 17–25.